

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2003, 20:11:57 ; Search time 7402 Seconds
(without alignments)
11712.680 Million cell updates/sec

Title: US-09-842-484A-1
Perfect score: 2979
Sequence: 1 ttataactgattaaagaag.....acattgcattttattaaaa 2979

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :					GenEmbl.*				
1:	gb_ba.*	2:	gb_htg.*	3:	gb_in.*	4:	gb_om.*	5:	gb_ov.*
6:	gb_pat.*	7:	gb_ph.*	8:	gb_pl.*	9:	gb_pr.*	10:	gb_ro.*
11:	gb_sts.*	12:	gb_sy.*	13:	gb_un.*	14:	gb_vi.*	15:	em_ba.*
16:	em_fun.*	17:	em_fun.*	18:	em_in.*	19:	em_mu.*	20:	em_om.*
21:	em_or.*	22:	em_ov.*	23:	em_pat.*	24:	em_ph.*	25:	em_pl.*
26:	em_ro.*	27:	em_sts.*	28:	em_un.*	29:	em_vi.*	30:	em_htg_hum.*
31:	em_htg_inv.*	32:	em_htg_other.*	33:	em_htg_mus.*	34:	em_htg_pln.*	35:	em_htg_rod.*
36:	em_htg_mam.*	37:	em_htg_vrt.*	38:	em_sy.*	39:	em_htgo_hum.*	40:	em_htgo_mus.*
41:	em_htgo_other.*								

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2979	100.0	2979	1	AF195517	Pasteurel
2	2635.8	88.5	11885	1	AE006116	Pasteurel
3	2632.6	88.4	8838	1	AF302467	Pasteurel
4	2340.8	78.6	16727	1	AF067175	Pasteurel
5	2283.4	76.6	2937	1	AF036004	Pasteurel
6	2268.6	76.2	2919	1	AF237926	Pasteurel
c 7	814.2	27.3	14483	1	AB079602	Escherich
c 8	97.6	3.3	155204	2	AC007926	Trypanoso
c 9	86.2	2.9	100726	2	AC116961	Dictyoste
c 10	82.6	2.8	43993	2	AC116965	Dictyoste
c 11	82.4	2.8	11442	1	AF400048	Campyloba
c 12	82.4	2.8	11455	1	AY044868	Campyloba
c 13	82.4	2.8	16265	1	VCLPSS	V.cholerae
c 14	82.4	2.8	46721	1	AB012956	Vibrio ch
c 15	82.4	2.8	270050	1	AL591977	Listeria
c 16	81.8	2.7	153026	9	AL391280	Human DNA
c 17	81.6	2.7	54345	3	AC084152	Caenorhab
c 18	81.6	2.7	302000	1	AP003187	Clostridi
c 19	81	2.7	984	6	AX413863	Sequence
c 20	81	2.7	984	6	AX415777	Sequence
c 21	80.8	2.7	12388	1	AF401529	Campyloba
c 22	80.8	2.7	12390	1	AF401528	Campyloba
c 23	80.8	2.7	15671	1	AF343914	Campyloba
c 24	80.6	2.7	9987	1	AB050723	Streptoco
c 25	80.6	2.7	62352	2	AC116990	Dictyoste
c 26	80.6	2.7	171798	2	AC084397	Trypanoso
c 27	80.4	2.7	18239	1	AF349539	Streptoco
c 28	80.4	2.7	86827	3	PFMAL3P5	Plasmodiu
c 29	80.4	2.7	347400	1	AP003591	Nostoc sp
c 30	80	2.7	19966	6	AF373595	Streptoco
c 31	80	2.7	19966	6	AX009404	Sequence
c 32	80	2.7	27782	1	AF448502	Streptoco
c 33	80	2.7	178273	2	AC005308	Plasmodiu
c 34	79.4	2.7	10827	1	AE006155	Pasteurel
c 35	79.2	2.7	840	8	CNS0180K	Botrytis
c 36	79.2	2.7	11474	1	AF130984	Campyloba
c 37	79.2	2.7	11474	1	AF167344	Campyloba
c 38	79.2	2.7	11474	1	AF215659	Campyloba
c 39	79	2.7	20416	3	AC114257	Dictyoste
c 40	79	2.7	107289	2	AC116923	Dictyoste
c 41	79	2.7	158548	3	PFMAL3P2	Plasmodiu
c 42	78.6	2.6	106434	3	AC117080	Dictyoste
c 43	78.2	2.6	12728	2	AC116101	Dictyoste
c 44	78	2.6	16563	2	AC115600	Dictyoste
c 45	78	2.6	104992	2	AC005504	Plasmodiu

ALIGNMENTS

RESULT 1
AF195517
LOCUS AF195517
DEFINITION Pasteurella multocida chondroitin synthase CS gene, complete cds.
ACCESSION AF195517
VERSION AF195517.1 GI:9716369
KEYWORDS
SOURCE Pasteurella multocida.
ORGANISM Pasteurella multocida
Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
REFERENCE 1 (bases 1 to 2979)
AUTHORS DeAngelis,P.L. and Padgett-McCue,A.J.
TITLE Identification and molecular cloning of a chondroitin synthase from Pasteurella multocida type F


```
QY 1381 GCTTAAACCTGTCGCAAAATTAATTTCAAGATGCTAGATAGTCTCTTAATCAAACTGTT 1440
|||||
Db 1381 GCTTAAACCTGTCGCAAAATTAATTTCAAGATGCTAGATAGTCTCTTAATCAAACTGTT 1440

QY 1441 GTCGATCTCGAGGTTGTGTTATTTGTAAGATGGTTCAACAGATAAATACCTTTAGAGTGATC 1500
|||||
Db 1441 GTCGATCTCGAGGTTGTGTTATTTGTAAGATGGTTCAACAGATAAATACCTTTAGAGTGATC 1500

QY 1501 AATAAGCTTTATGGTAAATATCTAGGTCAGCATCATGCTTAACCAAAATGGCGGAATA 1560
|||||
Db 1501 AATAAGCTTTATGGTAAATATCTAGGTCAGCATCATGCTTAACCAAAATGGCGGAATA 1560

QY 1561 GCCTCAGCATCAAAATGACGCGCTTCTTTGCTAAAGGTTATTACATTTGGCGAGTTAGAT 1620
|||||
Db 1561 GCCTCAGCATCAAAATGACGCGCTTCTTTGCTAAAGGTTATTACATTTGGCGAGTTAGAT 1620

QY 1621 TCAGATGATTATCTTGAGCGTGTAGTCAAGTTGAACTGTGTTTAAAGAAATTTTAAAGAT 1680
|||||
Db 1621 TCAGATGATTATCTTGAGCGTGTAGTCAAGTTGAACTGTGTTTAAAGAAATTTTAAAGAT 1680

QY 1681 AAAACGCTAGCTTGTCTTTATACCACTAATAGAAACGTCATCCGGATGCTTAATC 1740
|||||
Db 1681 AAAACGCTAGCTTGTCTTTATACCACTAATAGAAACGTCATCCGGATGCTTAATC 1740

QY 1741 GCTAATGGTTACAAATGGCCAGAAATTTTACGAGAAAACTCAACACGGCTATGATTGCT 1800
|||||
Db 1741 GCTAATGGTTACAAATGGCCAGAAATTTTACGAGAAAACTCAACACGGCTATGATTGCT 1800

QY 1801 CACCATTTTGAATGTTTACGATTAGAGCTTGGCATTTAACCGGATGATTAAACGAAAT 1860
|||||
Db 1801 CACCATTTTGAATGTTTACGATTAGAGCTTGGCATTTAACCGGATGATTAAACGAAAT 1860

QY 1861 ATTGAACCGCGTGGATTATGACATCTTCTTAACTCAGTGAAGTTGGAAAAATTTAAA 1920
|||||
Db 1861 ATTGAACCGCGTGGATTATGACATCTTCTTAACTCAGTGAAGTTGGAAAAATTTAAA 1920

QY 1921 CATCTTAAATAAATCTGCTATAACCGCGTATTACATGGTGATACACATCCATTAAGAAA 1980
|||||
Db 1921 CATCTTAAATAAATCTGCTATAACCGCGTATTACATGGTGATACACATCCATTAAGAAA 1980

QY 1981 CTGGCATTCAAAGAAAAACCATTTTGTGTAGTCAATCAGTCATTTAAATAGACAAGGC 2040
|||||
Db 1981 CTGGCATTCAAAGAAAAACCATTTTGTGTAGTCAATCAGTCATTTAAATAGACAAGGC 2040

QY 2041 ATCAATTTATTATATGACAAATTTTGATGATTTAGATGAAAGTAAAGTATATCTTC 2100
|||||
Db 2041 ATCAATTTATTATATGACAAATTTTGATGATTTAGATGAAAGTAAAGTATATCTTC 2100

QY 2101 AATAAAACCGCTGAATATCAAGAGAAATGGATATGTTTAAAGATCTTAAACTCATTTCAA 2160
|||||
Db 2101 AATAAAACCGCTGAATATCAAGAGAAATGGATATGTTTAAAGATCTTAAACTCATTTCAA 2160

QY 2161 AATAAGATGCCAAATCGCAGTCAGTATTTTCTATCCCAATACATTTAAACGGCTTAGTG 2220
|||||
Db 2161 AATAAGATGCCAAATCGCAGTCAGTATTTTCTATCCCAATACATTTAAACGGCTTAGTG 2220

QY 2221 AAAAACTAACCAATATTATGATATAATAAATAATATATCGTTATATCTACATGTT 2280
|||||
Db 2221 AAAAACTAACCAATATTATGATATAATAAATAATATATCGTTATATCTACATGTT 2280

QY 2281 GATAAGAAATCATCTTACACCAAGACATCAAAAAAGAAATATTGGCTTTCTATCATAAAGCAC 2340
|||||
Db 2281 GATAAGAAATCATCTTACACCAAGACATCAAAAAAGAAATATTGGCTTTCTATCATAAAGCAC 2340

QY 2341 CAAGTGAATATTTTACTAAATAATAGACATCTCATATTTACAGAGTAATAGACTAATAAAA 2400
|||||
Db 2341 CAAGTGAATATTTTACTAAATAATAGACATCTCATATTTACAGAGTAATAGACTAATAAAA 2400

QY 2401 ACTGAGGCACATTTAAGTAATATTAATAATTAAGTCAGTTAAATCTAAATTTGTGAATAC 2460
|||||
Db 2401 ACTGAGGCACATTTAAGTAATATTAATAATTAAGTCAGTTAAATCTAAATTTGTGAATAC 2460
```

```
QY 2461 ATCATTTTTGATAATATGACAGCCTATTCGTTAAAAATGACAGCTATGCTTATATGAAA 2520
|||||
Db 2461 ATCATTTTTGATAATATGACAGCCTATTCGTTAAAAATGACAGCTATGCTTATATGAAA 2520

QY 2521 AAATATGATGTGGCATGAATTTCTCAGCATTAACACATGATTGGATCGAGAAAATCAAT 2580
|||||
Db 2521 AAATATGATGTGGCATGAATTTCTCAGCATTAACACATGATTGGATCGAGAAAATCAAT 2580

QY 2581 GCGCATCCACCATTTAAAAAAGCTGATTAATAACCTATTTTAATGACAAATGACTTAAGAAGT 2640
|||||
Db 2581 GCGCATCCACCATTTAAAAAAGCTGATTAATAACCTATTTTAATGACAAATGACTTAAGAAGT 2640

QY 2641 ATGAATGTGAAGGGGCATCACAAGGTATGTTTATGAAGTATGGCTACCCGCATGAGCTT 2700
|||||
Db 2641 ATGAATGTGAAGGGGCATCACAAGGTATGTTTATGAAGTATGGCTACCCGCATGAGCTT 2700

QY 2701 CTGACGATTTAATAAGAGTCATCACATCTCGGCCAAATCAATGATAGTGCCAGAAATAT 2760
|||||
Db 2701 CTGACGATTTAATAAGAGTCATCACATCTCGGCCAAATCAATGATAGTGCCAGAAATAT 2760

QY 2761 AACACTGAGGATATTGGTTCCAAATTTGCCACTTTTAGAAAAAGAAAACCGGCCAT 2820
|||||
Db 2761 AACACTGAGGATATTGGTTCCAAATTTGCCACTTTTAGAAAAAGAAAACCGGCCAT 2820

QY 2821 GTATTTAATAAACATCGACCTGACTTATATATGCTTTGGGAACGAAAATTTACAATGGACA 2880
|||||
Db 2821 GTATTTAATAAACATCGACCTGACTTATATATGCTTTGGGAACGAAAATTTACAATGGACA 2880

QY 2881 AATGAACAAATTTCAAGTGTCAAAAAAGGCGAAAATATCCCGGTTTAAACAAGTTCAATT 2940
|||||
Db 2881 AATGAACAAATTTCAAGTGTCAAAAAAGGCGAAAATATCCCGGTTTAAACAAGTTCAATT 2940

QY 2941 AATAGTATAACGCTATAAAACATTTTGCAATTTTATATAAA 2979
|||||
Db 2941 AATAGTATAACGCTATAAAACATTTTGCAATTTTATATAAA 2979

RESULT 2
AE006116 11885 bp DNA linear BCT 08-MAR-2001
LOCUS Pasteurella multocida PM70 section 83 of 204 of the complete
DEFINITION genome.
ACCESSION AE006116 AE004439
VERSION AE006116.1 GI:12721075
KEYWORDS Pasteurella multocida.
SOURCE Pasteurella multocida.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
REFERENCE 1 (bases 1 to 11885)
AUTHORS May, B. J., Zhang, Q., Li, L. L., Paustian, M. L., Whittam, T. S. and
Kapur, V.
TITLE Complete genomic sequence of Pasteurella multocida, Pm70
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (6), 3460-3465 (2001)
MEDLINE 21145866
PUBMED 11248100
REFERENCE 2 (bases 1 to 11885)
AUTHORS Zhang, Q. and Kapur, V.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-2000) Department of Veterinary Pathobiology,
University of Minnesota, 1971 Commonwealth Ave., St. Paul, MN
55108, USA
FEATURES
source 1..11885
/organism="Pasteurella multocida"
/strain="PM70"
/db_xref="taxon:747"
205..2073
/gene="hyaE"
/notes="PM0774"
205..2073
/gene="hyaE"
/codon_start=1
gene
CDS
```


/product="unknown"									
/protein_id="AAK02866.1"									
/db_xref="GI:12721084"									
/translation="METLDKIKQISENPILYMKSPKFPSCGFSARAVEALMHCKV PGYVDILQHPDIRLELPAYANWTFPQLWVDGELVGGCDIILEMFQQGLQTLADV AAKYPQE"									
BASE COUNT	4287 a	2084 c	1962 g	3552 t					
ORIGIN									
Query Match 88.5%; Score 2635.8; DB 1; Length 11885;									
Best Local Similarity 93.0%; Pred. No. 0;									
Matches 271; Conservative 0; Mismatches 207; Indels 1; Gaps 1;									
Qy	1	TTATAACTGATTAAAGAAGGTAACGATTCAAGCAGGTTAATTTTAAAGGAAAGAAA	60						
Db	2031	TTATAAACTGATTAAAGAAGGTAACGATTCAAGCAGGTTAAATTTT - AAGGAAGAAA	2089						
Qy	61	ATGAATACATTATCACAAGCAATAAAAGCATATAACAGCAATGACTATGAATTAGCACTC	120						
Db	2090	ATGAATACATTATCACAAGCAATAAAAGCATATAACAGCAATGACTATGAATTAGCACTC	2149						
Qy	121	AAATTATTGAGAAGTCTGCTGAACCTACGGCGGAAAAATCGTTGAATTCCAAATATTC	180						
Db	2150	AAATTATTGAGAAGTCTGCTGAACCTACGGCGGAAAAATCGTTGAATTCCAAATATTC	2209						
Qy	181	AAATGTAAAGAAAACCTCTCGACCAATTTCTTATGTAAGTGAAGATAAAAAACAGTGT	240						
Db	2210	AAATGTAAAGAAAACCTCTCGACCAATTTCTTATGTAAGTGAAGATAAAAAACAGTGT	2269						
Qy	241	TGGCATAGCTCATTAGATATCGCAACACAGCTCTTACTTTCCACGCTAAAAAATTAAC	300						
Db	2270	TGGCATAGCTCATTAGATATCGCAACACAGCTCTTACTTTCCACGCTAAAAAATTAAC	2329						
Qy	301	CTATCCGAATCAGAAAAACAGTTTAAAAATAAATGGAATCTATCACTGGGAAAAAA	360						
Db	2330	CTATCCGAATCAGAAAAACAGTTTAAAAATAAATGGAATCTATCACTGGGAAAAAA	2389						
Qy	361	TCGGAGAACGCAAAATCAGAAAGGTGGAAGTCTAGTACCCAAAGATTTTCCATAAGATCTT	420						
Db	2390	TCGGAGAACGCAAAATCAGAAAGGTGGAAGTCTAGTACCCAAAGATTTTCCATAAGATCTT	2449						
Qy	421	GTCTTCTGCTCCATTGCCAGATCATGTTAATGATTTTACATGGTACAAAAATCGAAAAAA	480						
Db	2450	GTCTTCTGCTCCATTGCCAGATCATGTTAATGATTTTACATGGTACAAAAATCGAAAAAA	2509						
Qy	481	AGCTTAGGTATAAGCCTGTAAATAAGAAATATCGGTCTTTCTATTATTATTCCTACATTT	540						
Db	2510	AGCTTAGGTATAAGCCTGTAAATAAGAAATATCGGTCTTTCTATTATTATTCCTACATTT	2569						
Qy	541	AATCGTAGCCGTATTTTAGATATAACGTTAGCCCTGTTTGGTCAATCAGAAAAACAACTAC	600						
Db	2570	AATCGTAGCCGTATTTTAGATATAACGTTAGCCCTGTTTGGTCAATCAGAAAAACAACTAC	2629						
Qy	601	CCATTTGAAGTCGTTGTCAGATGATGGTAGTAAGGAAACCTTACTTACCATTTGTGCAA	660						
Db	2630	CCATTTGAAGTCGTTGTCAGATGATGGTAGTAAGGAAACCTTACTTACCATTTGTGCAA	2689						
Qy	661	AAATACGAACAAAAACTTTGACATAAAGTATGTAAGACAAAAAGATTATGGATATCAATTG	720						
Db	2690	AAATACGAACAAAAACTTTGACATAAAGTATGTAAGACAAAAAGATTATGGATATCAATTG	2749						
Qy	721	TGTGCACTCAGAACTTTAGGTTTACGTACAGCAAGTATGATTTTGTCTCGATTTCTAGAC	780						
Db	2750	TGTGCACTCAGAACTTTAGGTTTACGTACAGCAAGTATGATTTTGTCTCGATTTCTAGAC	2809						
Qy	781	TGGCATATGGCACACAACTATATGGGTTTCACTTTATCTTACAGAACTATTAGAAGAC	840						
Db	2810	TGGCATATGGCACACAACTATATGGGTTTCACTTTATCTTACAGAACTATTAGAAGAC	2859						
Qy	841	AATCATATTTGTTTAAATTTGGACCTAGAAAAATATGTGGATACTCATATATTTACCGCAGAA	900						
Db	2870	AATCATATTTGTTTAAATTTGGACCTAGAAAAATATGTGGATACTCATATATTTACCGCAGAA	2929						

QY	901	CAATTCCTTAACGATCCATATTTAATAGAATCACTACTGTAAGCCGCTACAAAATAACAAT	960						
DB	2930	CAATTCCTTAACGATCCATATTTAATAGAATCACTACTGTAAGCCGCTACAAAATAACAAT	2989						
QY	961	CTTTCGATTTACATCAAAAGGAATATATCTGTTGGATTGGAGATTAGAACATTTTCAAAAA	1020						
DB	2990	CTTTCGATTTACATCAAAAGGAATATATCTGTTGGATTGGAGATTAGAACATTTTCAAAAA	3049						
QY	1021	ACGGATAATCTACGTCATATGTTGATTCCTCGTTTCTGTTATTTTGGGGTAATGTTGCA	1080						
DB	3050	ACGGATAATCTACGTCATATGTTGATTCCTCGTTTCTGTTATTTTAGTTGGGGTAATGTTGCA	3109						
QY	1081	TTTTTCTAAAGAATGGCTAAATAAAGTAGTTGGTTTCGATGAAGAAATTTAATCATTTGGGG	1140						
DB	3110	TTTTTCTAAAGAATGGCTAAATAAAGTAGTTGGTTTCGATGAAGAAATTTAATCATTTGGGG	3169						
QY	1141	GGCGAAGATGTAGAAATTTGGTTACAGATTATTTGCCAAAGCTGTTTTTTCAGAGTAAT	1200						
DB	3170	GGCGAAGATGTAGAAATTTGGTTACAGATTATTTGCCAAAGCTGTTTTTTCAGAGTAAT	3229						
QY	1201	GACGGCGAATGGCCATCCATCAAGAACCCACTCTGTAAAGAAAAATGAAACAGAACGCGAA	1260						
DB	3230	GACGGCGAATGGCCATCCATCAAGAACCCACTCTGTAAAGAAAAATGAAACAGAACGCGAA	3289						
QY	1261	GCTGGTAAAAAGTATTACGCTTAAATTTGTGAAGAAAAAGGTACCTTACATCTATAGAAG	1320						
DB	3290	GCTGGTAAAAAGTATTACGCTTAAATTTGTGAAGAAAAAGGTACCTTACATCTATAGAAG	3349						
QY	1321	CTTTTACCAATAGAAATTCACATATTTCATAGAAATACCTTTAGTTTCTATTTTATATCCCC	1380						
DB	3350	CTTTTACCAATAGAAATTCACATATTTCATAGAAATACCTTTAGTTTCTATTTATATCCCC	3409						
QY	1381	GCTTTAATCTGCAAAATATATTCAAGATGTGTAGATAGTCTCTTAAATCAAACTGTT	1440						
DB	3410	GCTTTAATCTGCAAAATATATTCAAGATGTGTAGATAGTCTCTTAAATCAAACTGTT	3469						
QY	1441	GTGCATCTCGAGGTTGTATTGTGAACGATGGTTCAACAGATAATACCTTGAAGAGTATC	1500						
DB	3470	GTGCATCTCGAGGTTGTATTGTGAACGATGGTTCAACAGATAATACCTTGAAGAGTATC	3529						
QY	1501	AATAAGCTTTATGCTAATAATCCTTAGGTTAGGTCATGCTTAAACCAAAATGGCGGAATA	1560						
DB	3530	AATAAGCTTTATGCTAATAATCCTTAGGTTAGGTCATGCTTAAACCAAAATGGCGGAATA	3589						
QY	1561	GCCTCAGCATCAAAATGCAGCGGTTCTTTGCTAAAGGTTATTACATTTGGCGAGTTAGAT	1620						
DB	3590	GCCTCAGCATCAAAATGCAGCGGTTCTTTGCTAAAGGTTATTACATTTGGCGAGTTAGAT	3649						
QY	1621	TCAGATGATTTATCTTGAGCCCTGATGTCAGTTGAACTGTGTTTTAAAGAAATTTTTAAAGAT	1680						
DB	3650	TCAGATGATTTATCTTGAGCCCTGATGTCAGTTGAACTGTGTTTTAAAGAAATTTTTAAAGAT	3709						
QY	1681	AAAACGCTAGCTTCTGTTTATACCACATAATAGAAACGTCATCCGGATGGTAGCTTAATC	1740						
DB	3710	AAAACGCTAGCTTCTGTTTATACCACATAATAGAAACGTCATCCGGATGGTAGCTTAATC	3769						
QY	1741	GCTAATGTTTACAATTTGGCCAGAATTTTCAGAGAAAAAATCACAACGGCTATGATTTGCT	1800						
DB	3770	GCTAATGTTTACAATTTGGCCAGAATTTTCAGAGAAAAAATCACAACGGCTATGATTTGCT	3829						
QY	1801	CACCATTTTGAATTTTACGATTTAGAGCTTTGGCATTTTAACGGATGGATTTTAACGAAAT	1860						
DB	3830	CACCATTTTGAATTTTACGATTTAGAGCTTTGGCATTTTAACGGATGGATTTTAACGAAAT	3889						
QY	1861	ATTGAAAACGGCGTGGATTTATGACATGTTTCTTTAACTCAGTCAGTTGGAAAAATTTAAA	1920						
DB	3890	ATTGAAAACGGCGTGGATTTATGACATGTTTCTTTAACTCAGTCAGTTGGAAAAATTTAAA	3949						
QY	1921	CATCTTAATAAATCTGCTATAAACCCTATTACATGTTGATGATAACACATCCATTAAGAAA	1980						
DB	3950	CATCTTAATAAATCTGCTATAAACCCTATTACATGTTGATGATAACACATCCATTAAGAAA	4009						
QY	1981	CTCGGCATTTCAAAAGAAAAACCATTTTGTGTAGTCAATCAGTCATTTAAATAGACAAGGC	2040						

Db 4010 CTAGACACTCAAAAGAAAATCATTTTCGTTGCTTAATCAATCATTAATAATAGCGCAGAGA 4069
Qy 2041 ATCAATATTTATTAATATGACAAAATTTGATGATTTAGATGAAGTAGAAAGTATATCTTC 2100
Db 4070 GTAAGTAATTAACATATGATGAATTCGATAACTTACATTAATGATAAGTAGAAATATACATTTTC 4129
Qy 2101 AATAAAACGGTGAATATCAAGAAGAAATGGATATGTTTAAAAGATCTTAAACTCATTTCAA 2160
Db 4130 AATAAAACAGCAGATATCAAGAAGAAATGATATTTCTAAAAGATATCAAAAATAGTACAG 4189
Qy 2161 AATAAGATGCCAAATCGCAGTCAGTATTTCTTATCCCAATACATTAACCGGCTTAGTG 2220
Db 4190 CTTAAGATGCCAAAGTGGCAATAGTATATTTTATCCAAATAGATAGATGGGTTAGTT 4249
Qy 2221 AAAAACTAAACAATATTTATGAATATATAAATAATATATTCGTTATATTTACATGTT 2280
Db 4250 AAAAAATTAACAATATTTATCGAATATACAAAAATGTACTCATTTATGTTTACACATT 4309
Qy 2281 GATAGAATCATCTTACACCACATCAAAAAGAAATATTGGCTTCTCATCATAGCAC 2340
Db 4310 GACAAAAATCATCTCACTTACAGACATCAAAAAGGAATACTAGAATTTTCATACAGAAT 4369
Qy 2341 CAAGTGAATATTTACTAATAATGACATCTCATATTTACACAGTAAATAGACTTAATAAAA 2400
Db 4370 CAATCAATATCTTGTAAATATGATGTTCTTATTAACAACAACAGGCTCATAAAA 4429
Qy 2401 ACTGAGGCATTTAAGTAAATTAATAAATTAAGTCAGTTAAATCTAAATTTGTGAATAC 2460
Db 4430 ACTAAAGCTCATTTAAGCAATATGAACAAGTTAAGACAACCTAAATCTTAATTTAGAATAT 4489
Qy 2461 ATCATTTTGTAAATCATGACAGCCTATTCGTTAAATAAGTACAGCTATGCTTATATGAAA 2520
Db 4490 ATTATTTTGAACAACATGACAGCCTATTTATTAATAAATGATGTTTATACCATATAAAA 4549
Qy 2521 AAATATGATCGCGCATGAATTTCTCAGCATTAACACATGATTTGGATCGAGAAAAATCAAT 2580
Db 4550 AAATATGATCGCATGAATTTCTCCTCATTAACAAATGATTGGATTAATAAATAAATTAAT 4609
Qy 2581 GCGCATCCACATTTAAAAGCTGATTAACACCTATTTTAATCAATGACTTTAAGAAGT 2640
Db 4610 GCACACTCACCATTTAAAATCTGATAAAAAATATTTTAAATGATATGATTTAAAACA 4669
Qy 2641 ATGAATGTGAAGGCGCATCACAGGTATGTTTATGAAGTATGCGGTACCGCATGAGCTT 2700
Db 4670 ATAAATGAAGGCGCGCTCACAGGATGTTTCATAAATACACTTTAGCACATGATTT 4729
Qy 2701 CTGACGATTTATAAAGAGTCATCACATCTGCCAATCAATTTGATAGTGGCCAGATAT 2760
Db 4730 GCTACGATTTATGAAGAAGTCATTACATTTATGCCAATCTACAGATAGTGTCTGTAATAT 4789
Qy 2761 AACACTCAGATATTTGGTTCCAAATTTGCACTTTTAAATCTTAGAAAAAGAAACCGGCAT 2820
Db 4790 AATACTGAAGATATTTGGTTCCAAATTTGCACTTTTAAATCTTAGAAAAAGAAACCGGCAT 4849
Qy 2821 GTATTTATAAACAATCGACCTGACCTTATATGCTTTGGGAACGAAATTTACATGACA 2880
Db 4850 GTATTTATAAACAATCGACCTGACCTTATATGCTTTGGGAACGAAATTTACATGACA 4909
Qy 2881 AATGAACAATTTCAAGTGCNAAAAAAGGCAAAATATCCCGTTTACAAGTTCAATTATT 2940
Db 4910 AATGAACAATTTGAAGTGCNAAAAAAGGCAAAATATACCTGTTTACAAGTTCAATTATT 4969
Qy 2941 AATAGTATAAGCTATATAAATTTGCAATTTTATATAAA 2979
Db 4970 AATAGTATAACTCTATAAACAACATGCAATTTTATATAAA 5008

RESULT 3

AF302467

LOCUS

DEFINITION Pasteurella multocida P4218 region 2 capsule biosynthesis gene

cluster, partial sequence.

ACCESSION

VERSION GI:13274373

KEYWORDS

SOURCE

ORGANISM

Pasteurella multocida.
Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
Pasteurella.

REFERENCE

1 (bases 1 to 8838)
Townsend, K.M., Boyce, J.D., Chung, J.Y., Frost, A.J. and Adler, B.
Genetic organization of Pasteurella multocida cap Loci and
development of a multiplex capsular PCR typing system
J. Clin. Microbiol. 39 (3), 924-929 (2001)

JOURNAL

MEDLINE
21142635

PUBMED

11230405

REFERENCE

2 (bases 1 to 8838)
Townsend, K.M., Boyce, J.D., Chung, J.Y., Frost, A.J. and Adler, B.
Direct Submission
Submitted (03-SEP-2000) Veterinary Pathology and Anatomy, The
University of Queensland, School of Veterinary Science, Brisbane,
QLD 4073, Australia

FEATURES

source

1. 8838
/organism="Pasteurella multocida"

/strain="P4218"

/serotype="F:3"

/db_xref="taxon:747"

complement(<1..351)

/gene="phyA"

complement(<1..351)

/gene="phyA"

/note="similar to KpsS, LipA"

/codon_start=1

/transl_table=11

/product="PhyA"

/protein_id="AAK17919.1"

/db_xref="GI:13274374"

/translation="MSILYDIRLNMYOEFLIFSGKMLKIPYLSGFFTORLKMFSFV
TWKERTCILEWGYKASSRKARHFAQQHDLPTATIEDGLRSIGLGVGGYPPSPSLYVD
DIGIYDINQPSRL"

699..2567

/gene="fcbE"

699..2567

/gene="fcbE"

/note="similar to Hyae, KfiB"

/codon_start=1

/transl_table=11

/product="FcbE"

/protein_id="AAK17920.1"

/db_xref="GI:13274375"

/translation="MKKVIIGHKQSNQYDEKVFQCYGNMPLPSKREKMSPIEIGH
VLNKLPSLEHTPKNSVLSNKKSKIKKNSAKNSHKHAKTNTIOTSSIWNLSD
LMLANTEQNFQWSDPNAIQILDYWANLDPNIHFVYDKPENLFQYHSLLEALKLDK
HTVQEFQEWQTYNEKILTYFNKYKDRSVLLNTQOLQNTKSTLSEIYKHLSPDALY
YIYSEQLVNDADFALLAWDMIOKVNOYQHEKELELSTIKERQLEVFERYQLTEQ
KLSETQKEIQIKDENRKYKSERAKLTASVOSTSKILSEKEKETSICIKSENTKKEEK
IKDEAHHLTKTSLDSKEKALKTHODEIPALKIENENISVQEDMQEKFQENKROE
LEQELKAISSDKALLETENSKQTSVESLENENKYLALQLOLQOELEKLYIDQVLK
AKPLRYGAADRIKNTYRLGYKIQHRSLFGLIFLFFILFFYLLGFKREMKRYEYN
TLPPIHIEYEDAEHANRISKHSLYKLGVLFOEINNPFPKWLTPYKLIKKEGRFQGG"

2584..5481

/gene="fcbD"

2584..5481

/gene="fcbD"

/note="similar to PmcS"

/codon_start=1

/transl_table=11

/product="glycosyltransferase FcbD"

/protein_id="AAK17921.1"

/db_xref="GI:13274376"

/translation="MNTLSQAIKAYNSNDYELALKLFEKSAETYGRKIVRFQIIKCKE
KLSTNSVSDKNSVCDSDSLQTLSSNVKKLTLSSEKNSLKNKKKSIITGKKSE
NAEIRKVLVPKDFPKDLVLAFLPDHVNDFETWYKRRKSLGIRKPNKKNIGLSITPTF

Qy	1141	GGCGAAGATGTAGAAATTTGGTTTACAGATTATTTGCCAAAGCGTGTGTTTTTCAGAGTAATTT	1200
Db	3664	GGCGAAGATGTAGAAATTTGGTTTACAGATTATTTGCCAAAGCGTGTGTTTTTCAGAGTAATTT	3723
Qy	1201	GACGGCGGAATGGCCATCCATCAAGAACCACCTGGTTAAGAAATGAAACAGAACCGGAA	1260
Db	3724	GACGGCGGAATGGCCATCCATCAAGAACCACCTGGTTAAGAAATGAAACAGAACCGGAA	3783
Qy	1261	GCTGGTAAAGATATTACGCTTAAATTTGTGAAGAAAGGTACCTTACATCTATAGAAAG	1320
Db	3784	GCTGGTAAAGATATTACGCTTAAATTTGTGAAGAAAGGTACCTTACATCTATAGAAAG	3843
Qy	1321	CTTTTACC AATAGAGATTCACATATTCAATAGATACCTTTAGTTTCTATTATATCCCC	1380
Db	3844	CTTTTACC AATAGAGATTCACATATTCAATAGATACCTTTAGTTTCTATTATATCCCC	3903
Qy	1381	GCTTATACCTGTGCAAAATTTATTTCAAGATGTCTAGATAGTGTCTTAATCAAACTGTT	1440
Db	3904	GCTTATACCTGTGCAAAATTTATTTCAAGATGTCTAGATAGTGTCTTAATCAAACTGTT	3963
Qy	1441	GTGCGATCTCGAGGTTGTTTGTGAAGATGGTTTCAACAGATTAATACCTTAGAAGTGCAT	1500
Db	3964	GTGCGATCTCGAGGTTGTTTGTGAAGATGGTTTCAACAGATTAATACCTTAGAAGTGCAT	4023
Qy	1501	AATAAGCTTTATGGTAAATAAATCTTAGGTAGCGCATCATGTCTAAACCAAAATGGCGGAATA	1560
Db	4024	AATAAGCTTTATGGTAAATAAATCTTAGGTAGCGCATCATGTCTAAACCAAAATGGCGGAATA	4083
Qy	1561	GGCTCAGCATCAAAATGCAGCCGTTCTTTTGGCTAAAGGTTATTACATTTGGCGAGTTAGAT	1620
Db	4084	GGCTCAGCATCAAAATGCAGCCGTTCTTTTGGCTAAAGGTTATTACATTTGGCGAGTTAGAT	4143
Qy	1621	TCAGAGCATATCTTTCAGCGCTCATGCAGTTGAACGTGTGTTTAAAGAAATTTTAAAGAT	1680
Db	4144	TCAGAGCATATCTTTCAGCGCTCATGCAGTTGAACGTGTGTTTAAAGAAATTTTAAAGAT	4203
Qy	1681	AAAAACGCTAGCTGTGTTTATACCACTAATAGAAACGTCATTCGCGGATGGTAGCTTAATC	1740
Db	4204	AAAAACGCTAGCTGTGTTTATACCACTAATAGAAACGTCATTCGCGGATGGTAGCTTAATC	4263
Qy	1741	GCTTAATGGTTACAAATTTGGCCAGATTTTCCACGAGAAAACTCACAAACGGCTATGATTGCT	1800
Db	4264	GCTTAATGGTTACAAATTTGGCCAGATTTTCCACGAGAAAACTCACAAACGGCTATGATTGCT	4323
Qy	1801	CACCATTTTAGAATGTTTACGATTAGAGCTTGCGATTTAACCGGATGGATTTAACGAAAT	1860
Db	4324	CACCATTTTAGAATGTTTACGATTAGAGCTTGCGATTTAACCGGATGGATTTAACGAAAT	4383
Qy	1861	ATTGAAACCGCGTGGAATATGACATGTTCCCTTAAACTCAGTGAAGTTGGAATAATTTAA	1920
Db	4384	ATTGAAACCGCGTAGACTATGACATGTTCCCTTAAACTCAGTGAAGTTGGAATAATTTAA	4443
Qy	1921	CATCTTAATAAAATCTGCTATAACCGCGTATTACATGGTGATTAACACATCCATTAAGAAA	1980
Db	4444	CATCTTAATAAAATTTGTTTATACCGCGTATTACATGGTGATTAACACATCAATTAATAAAT	4503
Qy	1981	CTCGGCATTCAAAAGAAAAACCATTTTGTGTAGTCAATCAGTCAATTAATAGACAAGGC	2040
Db	4504	CTAGACACTCAAAAAGAAAAATCATTTGCTGTGTTAATCAATCAATTAATAAATAAAT	4563
Qy	2041	ATCAATATTATTAATATTATGCAAAATTTGATGATTTAGATGAAAGTAAAGTATATCTTC	2100
Db	4564	GTAAGTAATTAATACTATGATGATTCGATAACTTAGATGAAAGTAAAGTAAAGTATATCTTC	4623
Qy	2101	AATAAAACCGCTGAATATCAAGAAGAAATGGATATGTTTAAAGATCTTAAACTCATTTCAA	2160
Db	4624	AATAAAACAGCAGATATTATCAAGAAGAAATTTGATATTTCTTAAAGATATCAAAATAGTACAG	4683
Qy	2161	ATAAAGATGCCAAATTCGCGAGTCAGTATTTTCTATCCCAATACATTTAAACGGCTTAGTG	2220
Db	4684	CGTAAGATGCAAAAGTTGCAATAGTATATTTTATCCAAATAGATAGATGGGTTAGTT	4743
Qy	2221	AAAAAACTAAACATATTATTTTGAATATTAATAAAAAATATATTTCGTTTATTTTTCATCATGTT	2280

Db	4744	AAAAATTAACAATATTATCGAATATAACAAAAATGTACTCATTTGTTTTACACATT	4803
Qy	2281	GATTAAGAAATCATCTTACACGAGACATCAAAAAGAAATATTGGCTTCTTATCATAGCAC	2340
Db	4804	GACAAAATCATCTCATCTTCAGACATCAAAAAGGAAATACTAGAATTTCATATAACAAGAT	4863
Qy	2341	CAAGTGAATATTTTACTAAATAAGTACATCTCATATTCACGAGTAAATAGACTAATAAAA	2400
Db	4864	CAATCAATATCTTGTTTAAATAATCATGTCTTCTTATATACAAAACACAGGCTCATAAA	4923
Qy	2401	ACTGAGGCACATTTTAAGTAATATTAATAAATTAAGTCAGTTAAATCTAAATTTGTGAATAC	2460
Db	4924	ACTAAGGCTCATTTTAAGCAATATGACACAGTTTAAGCACTAAATCTTAATTTAGAATAT	4983
Qy	2461	ATCATTTTTGATAATCATGACAGCCTATTTCGTTAAAAATGACAGCTATGCTTATATGAAA	2520
Db	4984	ATTATTTTGACAACCATGACAGCCTATTTTATTAATAATGATAGTTATACCCATATAAAA	5043
Qy	2521	AAATATGATGTCGGCATGAATTTCTCAGCATTAACACATGATTGGATCGAGAAAAATCAAT	2580
Db	5044	AAATATGATATCGGCATGAACTTCTCCTCATTAACAAATGATTTGGATATAATAAATTAAT	5103
Qy	2581	GGGCATCCACCATTTAAAAAGCTGATTAACACCTATTTTAATGACAACTGACTTTAAGAAGT	2640
Db	5104	GCACACTCACCATTTTAAAAATCTGATAAAAAAATATTTTAAATGATGATTTTAAAAACA	5163
Qy	2641	ATGAATGTGAAGGGCCATCACANGTATGTTTATGAAGTATCGCGTACCGCATGAGCTT	2700
Db	5164	ATAAATATGAAGGGCGGTACAAAGGAATGTTTCATTAATACACTTTAGCACATGATATT	5223
Qy	2701	CTGACGATTATTAAAGAAAGTCATCACATCCTGCCAATCAATGATAGTGTGCGCAGAATAT	2760
Db	5224	GCTACGATTATGAAGAAGTCATTACATTTATGCCAATCTACAGATAGTGTCTCTGAATAT	5283
Qy	2761	AACACTGAGGATATTGGTTCCAAATTTGCACCTTTTAATCTTTAGAAAAAGAAAACCGGCAT	2820
Db	5284	AATACTGAAGATATTGGTTCCAAATTTGCACCTTTTAATCTTTAGAAAAAGAAAACCGGCAT	5343
Qy	2821	GTATTTTAATAAAACATCGACCCCTGACTTATATGCTTTGGGAACGAAATTTACAATGGACA	2880
Db	5344	GTATTTTAATAAAACATCGACCCCTGACTTATATGCTTTGGGAACGAAATTTACAATGGACA	5403
Qy	2881	AATGAACAAATTTCAAATGCAAAAAAAGGGGAAAATATCCCGGTAAACAAGTTCATTATT	2940
Db	5404	AATGAACAAATTTGAAGTGTGAAAAAGAGGAGGAAAATATACCTGTTACAAGTTCATTATT	5466
Qy	2941	AATAGTATAACGCTATAAAACATTTGCAATTTTATTAATAA	2979
Db	5464	AATAGTATAACCTTATAAAACACTTGCATTTTATTAATAA	5502

RESULT 4	
AF067175	16727 bp DNA linear BCT 05-MAY-2000
LOCUS	Pasteurella multocida capsule biosynthesis gene cluster, complete
DEFINITION	sequence.
ACCESSION	AF067175
VERSION	AF067175.2 GI:7710189
KEYWORDS	.
SOURCE	Pasteurella multocida.
ORGANISM	Pasteurella multocida Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; Pasteurella. 1 (bases 1 to 16727) Chung, J.Y., Zhang, Y. and Adler, B. The capsule biosynthetic locus of Pasteurella multocida A:1. JOURNAL FEMS Microbiol. Lett. 166 (2), 289-296 (1998)
REFERENCE	98442423
AUTHORS	9770287
TITLE	2 (bases 1 to 16727)
JOURNAL	Chung, J.Y., Zhang, Y. and Adler, B.
MEDLINE	Direct Submission
PUBMED	
REFERENCE	
AUTHORS	
TITLE	

JOURNAL

Submitted (19-MAY-1998) Microbiology, Monash University, Wellington
Road, Clayton, VIC 3168, Australia
3 (bases 1 to 16727)
Chung, J. Y., Zhang, Y., and Adler, B.
Direct Submission

REFERENCE

Submitted (05-MAY-2000) Microbiology, Monash University, Wellington
Road, Clayton, VIC 3168, Australia
Sequence update by submitter

On May 5, 2000 this sequence version replaced gi:3435183.

REMARK

COMMENT

FEATURES

source

1..16727

/organism="Pasteurella multocida"

/db_xref="taxon:747"

/complement(1157..2383)

/gene="phyB"

/complement(1157..2383)

/gene="phyB"

/note="LipB, kps homolog"

/codon_start=1

/transl_table=11

/product="PhyB"

/protein_id="AAC67247.1"

/db_xref="GI:3435184"

/translation="MDVLKHNLDLLESNTYLLLGQPIGPFTEISAWLTKHKQVF
KINFSGDEAFYNSIPOTFAVQGLDQYLRFFCHQHSIDAIICFGDTRAYHRIA
KIVADQKISFWFEFGYLRPEYVTELEKGVNDYISOLPHANFELISQAACCSISAP
KLAKGFIPLATIAIKYTKAYFNQDKYPDYQHNRILDLKYIKLWLTSGTKRAYYQW
DRSFGKVAAGKGLDFELPQVDDSOVRIHCDYESDAFLKEILSFTHAPSALN
LIKKHPMRGFDYGHVTKYQTPAKKRIFYVHDVPTILLRKGKMITLNSYS
GIALLRMPVLVLGRASYNFEGLTGYNLHSEFHNKGLPEAKVFAYRYKYLKSTQI
NGSFYSKVLIRYPYQ"

/complement(2393..4483)

/gene="phyA"

/complement(2393..4483)

/gene="phyA"

/note="KpsS, LipA homolog"

/codon_start=1

/transl_table=11

/product="PhyA"

/protein_id="AAC67248.1"

/db_xref="GI:3435185"

/translation="MSTLYDIIRLNYQEFILPSKQMLKTYPLSGFRTORLKMFSPEV
TWKRTCLIEWGYSKSKRAHFAQHDPYATIEDFLRSIGSLGVDGYPPFSLVYD
DIGIYDIQPSLENLILSQDLQKVDQVEYATIELICTHNLKYNHAIPTPLONT
KRPIVLVDQYTGMAVTGNAEQSDFLHMLERAIITENPTAEILKTHPDMVCGKQOG
YLTEYQQFPRVKVLSDEPNSISLLKHVDKVCYVTSHTGFEALLLGTIVYTFGAWSG
WGLTDRHAYIRQKSKRAKRLQLQFATAFYQCRYINPNTGSGTDFDVIDLI
QAKKVTNLAGDIYCGMRPFWRKRVQPPFPFRCRLHFVNLHELKRCITHEKAQAKI
VYMGSHIEVVEYAKOQLPLLRMEDGFLRSVGLGSLNLTPTISLVLDVDGVIYFDAQSR
SRLEDILQHSFTLKDLQRAETLTKTLIEQHIGKYNVGHPTHLCLTHRONKLLVVGQV
ENDASIQGSPHRTNAELLCVTRKNNPOAYIYKPHDPVAGNRKNTDRLDYROYA
DFVVEKVNILDCINQVDEVTMTSLAGFEALLREKKVHCYGLFPYSNWGLTVDHLSLN
RRSRKLSLELIAGVLIYTPQYIDPTKTMDVQRAVDIILIEQRRIKNNKLNHTNFM
NIFMKLNLYSVLR"

4832..6700

/gene="hyae"

4832..6700

/gene="hyae"

/note="KfiB homolog"

/codon_start=1

/transl_table=11

/product="Hyae"

/protein_id="AAC67249.1"

/db_xref="GI:3435186"

/translation="MKKVIILGHKQSNYQVEKVCQYGMNPLPDKREKMSPIETGH
VLNKVLPSFTEHTPKNVLSLNSKSIKKGSAKNSKHKHAKTNTIOTTSIWNLNSD
LMLANTEQWGSNDPNAIQILDYWANLPNIHFVYIDKPNFLFYHLSLEALKDLK
HTVQERFEWQTYNEKILTYFNKYKDRSVLLNTQOLQNTKKTSLSEIYKHISAPDALY
KLINEPLNKEMEIENVDLSHQECPLSNFIVSOLIKNSPTVQVYVELOSHADLP
YISEQNLANDPALLAWKDMIOQYDANOYHEKELELSTIKEROLEVTEYKQLTEQ
KUSEQNLQEQIKDENRKKVSEKALFASVQSTKILSEKEKICISKSNTKIKEK
IKIDEAYHTKTLSDKALKRTHODEIALKIIFNENTSVQEDMQEKFQETNKKQOE
LEQELKAISDKALLETENSQKTQVSESLENENKVLLAQLQLIQEELEKLYIDNQVLK

gene

CDS

AKPRLYGAAIRIKNLQTYRLGYKIQRHGRSLFGLIFLFPFLFTYTLGFKREMKMYENN
TLPPIHVEYDAHEANRIKSHLSYKGLVFLQEIINNPEKWLTPYKLIKEGRRKQG"

/gene="hyAD"

6717..9635

/gene="hyAD"

/note="HyAD"

/codon_start=1

/transl_table=11

/product="glycosyl transferase"

/protein_id="AAC67250.2"

/db_xref="GI:7710190"

/translation="MNTLSQAIRAYNSNDYELALKFEKSAEYGRKIVFQITCKRE
KLSAHPSVNSAHLNVKKEKVNCDSPDLATOLLNVKKLVSLSDEKNTLKNKWL
LTEKKSNSAHRVALVPKDPKDLVLAPLPDHVDFTWYKKRKLGIKPEHQHVG
SIIVTFNRPAIISITLACLINQKTHYFVIVTDDGSEDLSPIIRQYENKLDIRYV
KQKNGFQASAARNMGLAKYDFIGLDCDMAPNPLVSHVSAELLEDLDTLIGPR
KYIDQHDIPKDFLNNASLESLEPVTNNNSAAKGGTWSLMDRQFQEFKTNLRIS
DSPFRFAAGNAFAKAKLNKSGFDEEFNHMGDEDFYGLFRYSGSFFKTIQINA
YHOEPPKENETDREACKNTLDIMREKVPYIRKLLP IEDSHNRVPLVSIYIPAYN
CANYIORCVDSALNOTVVDLEVCICNDGSTDNTLEVINKLYGNPRVIRINSPGSLA
SASNAVSFAKGYIIGQLDSDDDYLEPDAVELCLKEFLKDKTLACVYTNRNVPDGLS
IANGYNMPFSREKLTMTAMIAHHPMETIRAWHUTDGFNEKIENAVDTDFMLKLEVG
KFHLNKKICYNRVLHGDMTSSIKLGIOQKNHVVYVNSLNRRQGTYYNYDEFDLDES
RKYIFNKATYQEEIDLDKIKIIONKDAKIAVIFYPNTLNGVLKLNNTIENKNI
FVTVLHVKNHLTPDIKKELAPYKHQVNLINDISYVTSNRLIKTEAHLNINKL
SOLNLCXEYIIFDNHDSLVKNDSYAYMKYDVGWMSFALTDHWEIKINAHPPFKLI
KTPVNDNLSMNVKGASQGMFTYALAHALLTTIKEVITSCSIDSPYENTEDIWF
QFALLILLEKTHGVNKTSTLTUTYMPWERKQLMTWEQIESAKRGENIPVKNKIISITL

9699..10871

/gene="hyac"

9699..10871

/gene="hyac"

/note="hyac"

/codon_start=1

/transl_table=11

/product="UDP-glucose dehydrogenase"

/protein_id="AAC67251.2"

/db_xref="GI:7710191"

/translation="MKKITTAGCYGSLSNVLAQHNNVILLDIDQNKVDLINKKKS
PITDKEIEDLQNKSLMTMATDKVAKLNADFVIAITPTDYNTEGYFNTSTVEAVI
EGTSLINQATIIKISTIPVGFTEMRKFNTPNLIFSPFVRECKALYDNLVPSRII
VQTSYQAKYFADMLTQCARKQDVTLFTHTTEAEAVKLFANTYLANRVAPFNELDTY
ASLHLNKTDOIINGISDPRIGTHYNNPSGCGYCLPKDTKOLLANVADVPQNLIEA
IVKSNTRKRFEITHDVLNKPXTVGIYIRLINKSGSDNFRASAILDIMPFLKENGVEIV
IYEPTLNQAFEDYIPVNLQSEFINRSDDVILANKSEPDNLNOCSHKIYTRDIFGGDA"
10864..12297

/note="glycosyl transferase"

/codon_start=1

/transl_table=11

/product="HyAB/Hyaa fusion protein"

/protein_id="AAC67252.2"

/db_xref="GI:7710192"

/translation="MLNLFKIKRYVHIIQSLHKKKEYALLVAKYINLLISNQAYVIC
OLKLYDLIDPKWHSVFFQGLIARGHADNDVMVRLITCTDFSKNKLILSOLL
AYSPQIATTPKTYRYRALSIVANLKDPVRLKEELNKLPSCLVNTKPHYCLLQNE
VERENSKKLENIQOFLFYKLGTEITTIKTEPPNTNIVSHVSNNDNTYINARLITI
LVTFNSQKSIKNTLSNFOSYPNIEIIVIDDSQDNTSILQAYTKQYKNKIISL
KENYGVYAKNIGLKYASGEFITCODSDWAHPQKLALQVAPLLOHRELIVTFSKWVR
LDIPGNPYARTIYPLMRINPSALLFRKEKVECTALMDWVRIGADSEFNARLKLIFGH
KGYTVNKNPLTFGAHRENSLMTAQSTGYVNGVSLPREAYKNKWNHWHISQLOQRHAPL
LSSSPKREFQYDQHPIMDVISYAIN"

12316..13497

/gene="hexD"

12316..13497

/gene="hexD"

/note="CpxD, BexD homolog"

/codon_start=1

/transl_table=11

/product="HexD"

/protein_id="AAC67254.1"

/db_xref="GI:3435191"

gene

CDS

/translation="MKPIKTLISFITTLLIAGCHSMPTSGPAQNHILIGLKKPONESL
PSVDIEMNDKVAITLFEKQOSQFTQKQNSNYADIINVGDITLDVLWEAPAILF
GSLVSQTSKGNANLITPEQIVARNKGIITIPFGLVILVKGTPEQIORDLAHALSSLA
NKPOVIVRLNKNKSNVTILLRQNSVRMPLTSGQERVLDAIAVGGATENLDLSVOL
TRGEVKLSLEKLATHPENILLRNSDVITLKLPLSFTGLGALGNTKVKFSANGL
TLABGICEMGLLDNRADPKGVFVRIYIPFNKLSLAQETKWKARGYDNMEIPTVSV
LLNPNALFWLQRFPIQDKDLVYVSNAPMAEFQKFLKLVFSITSPVTGTLHNINIKN
L"
13507..14643
/gene="hexc"
CDS
13507..14643
/gene="hexc"
/note="BexC, CpxC homolog;
MIETKIQKNOFKKWSRFRKISTLFKVMVILPTCCSLFYFSLWASDIYISQSPFVVT
PKNGAALSGVALLOSGFARAQDDITYVQEFMRSTLLELKSIRPIQFVEDKGD
FSRNPLISEQAFIYQFSKLSVNFDSVSGIATNIRAFDPKRAQQINOELLKOG
EYLINLRNARKDITVFAELAVSEAKKVTETSSALSEYRKINGVFDLOSQSEVLS
LISTQORLVIQTLQDQVRSISPNPQVOTLLARANSIRKEMQOQVQVILGGNSIV
TOTAEYORLVNNTLAQQLGTAITSQNARSEADROQLYLEIISYPNEPDLAEPYR
LYNTLATFLSILYGITLLLASIREHN"
/codon_start=1
/transl_table=11
/product="HexC"
/protein_id="AAC67255.1"
/db_xref="GI:3435192"
/translation="MIETKIQKNOFKKWSRFRKISTLFKVMVILPTCCSLFYFSLWA
SDIYISQSPFVVTPKNGAALSGVALLOSGFARAQDDITYVQEFMRSTLLELLEK
SIPRIQFVEDKGDIFSERNPLNISEQAFIYQFSKLSVNFDSVSGIATNIRAFDP
KEAQINOELLKOGYEYLINLRNARKDITVFAELAVSEAKKVTETSSALSEYRKIN
GVFDLOSQSEVLSLISTQORLVIQTLQDQVRSISPNPQVOTLLARANSIRKEMQ
OQVQVILGGNSIVTOTAEYORLVNNTLAQQLGTAITSQNARSEADROQLYLEIIS
YYPNEPDLAEPYRLYNTLATFLSILYGITLLLASIREHN"
14643..15440
/gene="hexB"
CDS
14643..15440

Query Match 78.6%; Score 2340.8; DB 1; Length 16727;
Best Local Similarity 86.9%; Pred. No. 0;
Matches 2606; Conservative 0; Mismatches 372; Indels 22; Gaps 2;
QY 1 TTATAACTGATTAAAGAAGGTAACGATTCAAGCAAGGTTAAATTTTAAAGGAAGAAA 60
Db 6658 TTATAACTGATTAAAGAAGGTAACGATTCAAGCAAGGTTAAATTTT-AAAGGAAGAAA 6716
QY 61 ATGAATACATTTACACAGCAATAAAGCATATACAGCAATGACTATGAATTAGCACTC 120
Db 6717 ATGAATACATTTACACAGCAATAAAGCATATACAGCAATGACTATGAATTAGCACTC 6776
QY 121 AAATTATTTGAGAAGTCTGCTGAAACCTACGGCGGAAAAAATCGTTGAAATCCAAATATATC 180
Db 6777 AAATTATTTGAGAAGTCTGCGGAAATCTATGGACGAAAAATTTGTTGAATTTCAAAATTACC 6836
QY 181 AAATGTAAGAAAACTCTCGACCAATTC-----TTATGTTAACT 219
Db 6837 AAATGTAAGAAAACTCTCAGCACATCTCTGTTAAATTCAGCACATCTTTCTGTGTAAT 6896
QY 220 GAAGATAAAAAAACAAGTCTTTCGGATAGCTCATATAGATATCGCAACACAGCTCTTACTTT 279
Db 6897 AAAGAAGAAAAAGTCAATGTTTTCGGATAGCTCGGTAGATATTCACACACAACTGTACTTT 6956
QY 280 TCCACGTAATAAAAAATTAATCTATCCGAATCAGAAAAAACAAGTTTAAAAAATAAATGG 339
Db 6957 TCCACGTAATAAAAAATTAATCTATCTGACTCGGAAAAAACAAGTTTAAAAAATAAATGG 7016
QY 340 AAATCTATCACTGGGAAAAAATTCGAGAGACGGAGAAATCAGAAAGTGGAACTAGTACCC 399
Db 7017 AAATCTCACTGAGAGAAATTCGAGAAATTCGAGAGTGAAGACGGCTGCCCTTGTAACA 7076
QY 400 AAAGATTTTCCCTAAACATCTGTGTTCTGCTCCATCCAGATCATCTTAATGATTTTACA 459
Db 7077 AAAGATTTTCCCAAGATCTGGTTTATAGCCCTTTACCTGATCATCTTAATGATTTTACA 7136
QY 460 TGTACAAAAATCGAAAAAAGCTTTAGGTATATAAGCCCTGTAAATAAGAAATATCGGTCTT 519

Db 7137 TGGTACAAAAAGCGAAAAAGAAAGACTTGGCATAAAAACCTGAACATCAACATGTTGTCTT 7196
QY 520 TCTATTATTATTCCTACATTTTAATCGTAGCGGTATTTTAGATATATAACGTTAGCCTGTG 579
Db 7197 TCTATTATCGTTTACACAAATTCATCGTACCGCAATTTTATCGATTACATTTAGCCTGTGTTA 7256
QY 580 GTCATCAGAAAAAACAATACCCATTTTGAAGTCGTTTGGCAGATGATGGTAGTAGAGAA 639
Db 7257 GTAAACCAAAAAACAATACCCGTTTGAAGTTATCGTGACAGATGATGGTAGTAGAGAA 7316
QY 640 AACTTTACTTACCATTTGTCAAAAAATACGAACAAAACTTTGACATATAAGATATGTAGACAA 699
Db 7317 GATCTATCACCGATCATTCGCGCAATATGAAAAATAAATTTGGATTTCGCTACGTCAGACAA 7376
QY 700 AAGATTATGGATATCAATTTGTGTCAGTCAGAAAACTTAGGTTTACGTACACACAAAGTAT 759
Db 7377 AAGATAACGGTTTTTCAAGCCAGTCGCGGTGCGGAATATGGGATTTAGCTTACACAAATAT 7436
QY 760 GATTTTGTCTCGATTCTTAGACTGCGATATGGCACCACACAATTTATGGTTTCATCTTAT 819
Db 7437 GACTTTATTGGCTTACTCGACTGTGATATGGCGCAATCCATTTATGGTTTCATCTTAT 7496
QY 820 CTTACAGAACTATTAGACAGCAATGATATTGTTTAAITTTGGACCTTAGAAAAATATGTGGAT 879
Db 7497 GTTCAGAGCTATTATAGAGATGATGATTTAACAATCATTTGGTCCAGAAAAATACATCGAT 7556
QY 880 ACTCATATATTACCGCAGAACAAATTCCTTAACGATCCATATTTAATAGAATCAGTACCT 939
Db 7557 ACACAACTATTGACCCCAAAAGACTTCTTAAATACGCGAGTTTGGTTGAATCATTTACCA 7616
QY 940 GAAACGGGTACAAAAATAACAATCCTTCGATTACATCAAAAGGAAATAATATCGTTGGATTGG 999
Db 7617 GAAGTGAACCAACATAATAGTGTGCCGCAAAAGGGGAAGAACAGTCTTCTCTGGATTGG 7676
QY 1000 AGATTAGAACATTTCAAAAAACCGATATCTACGCTCTATGTGATTCCTCGTTCTGTTAT 1059
Db 7677 CGCTTAGAACAAATTCGAAAAACACAGAAAAATCTCCGCTTATCCGATTTCGGCTTTCCGTTT 7736
QY 1060 TTTGTTTCGGGGTAATGTTTCATTTTAAAGAAATGCTTAAATAAAGTAGTGTGTTTCGAT 1119
Db 7737 TTTGGCGGGTAATGTTTCGCTTTCGCTTAAATAATGCTTAAATAAATCCGTTTCTTTGAT 7796
QY 1120 GAAGAAATTTAATCATTTGGGGGCGGAAGATGTAGAATTTGGTTTACAGATTATTTGGCCAAA 1179
Db 7797 GAGGAATTCATCACTCGGGTGGAGAAGATGTGGAATTTGGATATCGCTTATTCCTTAC 7856
QY 1180 GSGCTTTTTCAGAGTAAATGACGCGGAATGGCCATCCATCAAGAACCCACCTGCTGATAA 1239
Db 7857 GGTAGTTTCTTTAAACTATTGATGGCATATGGCCCTACCATCAGAGCCACCGGTAAA 7916
QY 1240 GAAATGAACAGACGCGGAAGCTGGTAAAGATTTACGCTTAAATAATTTGTAAGAAAAAG 1299
Db 7917 GAAATGAACCGATCGTGAAGGGGAAAAAATATTACGCTCGATATATTAGAGAAAAAG 7976
QY 1300 GTACCTTACATCTATAGAAGCTTTTACCAATAGAAGATTCACATATTCATAGATACT 1359
Db 7977 GTCCCTTATATCTATAGAAAACTTTTACCAATAGAAGATTCGATATCAATAGATACCT 8036
QY 1360 TTAGTTTCTTATTTATATCCCGCTTAACTGTGCAAAATATATTTCAAAAGATGTTAGAT 1419
Db 8037 TTAGTTTCTTATTTATATCCAGCTTATTAAGTGTGCAAACTATATTTCAACGTTCCGTAGAT 8096
QY 1420 AGTGTCTTTAATCAAACTGTTGTGATCTCGAGGTTTGTATTTTGAACGATGTTTCAACA 1479
Db 8097 AGTGCACTGAATCAGACTGTTGTTGATCTCGAGGTTTGTATTTTGAACGATGTTTCAACA 8156
QY 1480 GATAATACCTTAGAAGTATCAATAAGCTTTATGGTAATAATCCTAGGGTACGATCATG 1539
Db 8157 GATAATACCTTAGAAGTATCAATAAGCTTTATGGTAATAATCCTAGGGTACGATCATG 8216
QY 1540 TCTAAACCAAAATGCGGAAATAGCCTCAGCATCAAAATGACGCGCTTTCTTTTGTCAAGGT 1599

Db 8217 TCTAAACCAATGCGGAATAGCCTCAGCATCAAAATGAGCCGTTTCTTTGCTAAAGGT 8276

Qy 1600 TATTACATTGGGCAGTTAGATTTCAGATGATTTATCTTGAGCCGTGATGCGATTGAACGTGTGT 1659

Db 8277 TATTACATTGGGCAGTTAGATTTCAGATGATTTATCTTGAGCCGTGATGCGATTGAACGTGTGT 8336

Qy 1660 TTTAAAGAAATTTTAAAGATATAAACGCTAGCTGTGTGTTTATACCACCTAATAGAAACGTC 1719

Db 8337 TTAAGAGATTTTAAAGATATAAACGCTAGCTGTGTGTTTATACCACCTAATAGAAACGTC 8396

Qy 1720 AATCCGATGATGCTTAATCGCTTAATGTTTACAAATTTGGCCAGAAATTTTTCAGAGAAAA 1779

Db 8397 AATCCGATGATGCTTAATCGCTTAATGTTTACAAATTTGGCCAGAAATTTTTCAGAGAAAA 8456

Qy 1780 CTCACAAGGCTATGATTGCTCACCATTTTAGAAATTTTACGATTAGAGCTTGGCATTTA 1839

Db 8457 CTCACAAGGCTATGATTGCTCACCATTTTAGAAATTTTACGATTAGAGCTTGGCATTTA 8516

Qy 1840 ACGGATGGATTTAAACGAAATATTGAAACGCCGTGGATTATGACATGCTTCCCTTAAACCTC 1899

Db 8517 ACTGATGGATTCATGAAGAAATTTGAAATGCCGTAGACTATGACATGCTTCCCTCAACCTC 8576

Qy 1900 AGTGAAGTTGGAATTTTAAACATCTTAAATAAATCTGCTATAACCGCGTATTACATGGT 1959

Db 8577 AGTGAAGTTGGAATTTTAAACATCTTAAATAAATCTGCTATAACCGGTATTACATGGT 8636

Qy 1960 GATACACATCCATTAGAAACTCGGCATTCAAAAGAAAACCATTTTGTGTAGTCAAT 2019

Db 8637 GATACACATCCATTAGAAACTTTGGCATTCAAAAGAAAACCATTTTGTGTAGTCAAT 8696

Qy 2020 CAGTCATTAAATAGACAAGGCATCAATTTATTATTAATATGACAAATTTGATGATTAGAT 2079

Db 8697 CAGTCATTAAATAGACAAGGCATCAATTTATTATTAATATGACAAATTTGATGATTAGAT 8756

Qy 2080 GAAAGTAGAAGTATATCTTCAATAAACCCGCTGAATATCAAGAGAAATGGATGTGTA 2139

Db 8757 GAAAGTAGAAGTATATCTTCAATAAACCCGCTGAATATCAAGAGAGATTTGATATCTTA 8816

Qy 2140 AAGATCTTAAACATCAATCAAAATAGAGATGCCAAATCGCAGTCAGTATTTTCTATCCC 2199

Db 8817 AAGATCTTAAACATCAATCAAAATAGAGATGCCAAATCGCAGTCAGTATTTTCTATCCC 8876

Qy 2200 AATACATTAAACGGCTTAGTGAAAAAAGCTAAACAATATTATTGAATATAATAAAAAATATA 2259

Db 8877 AATACATTAAACGGCTTAGTGAAAAAAGCTAAACAATATTATTGAATATAATAAAAAATATA 8936

Qy 2260 TTCGTTATTATCTACATGTTGTAAGAATCATCTTACACAGACATCAAAAAAGAAATA 2319

Db 8937 TTCGTTATTATCTACATGTTGTAAGAATCATCTTACACAGACATCAAAAAAGAAATA 8996

Qy 2320 TTGGCTTCTATCATAGCACCAGTGAATATTTTACTTAATATGACATCTCATATTAC 2379

Db 8997 CTAGCCTTCTATCATCAACATCAAGTGAATATTTTACTTAAATAATGATATCTCATATTAC 9056

Qy 2380 ACGAGTAATAGACTAATAAAAAAGCTGAGGCACATTTAAGTAATATTAATAAATTAAGTCAG 2439

Db 9057 ACGAGTAATAGACTAATAAAAAAGCTGAGGCACATTTAAGTAATATTAATAAATTAAGTCAG 9116

Qy 2440 TTAATCTAAATGTTGAATACATCATTTTGTGAATCATGACAGCCTATTCGTTAAAAAT 2499

Db 9117 TTAATCTAAATGTTGAATACATCATTTTGTGAATCATGACAGCCTATTCGTTAAAAAT 9176

Qy 2500 GACAGCTATGCTTATATGAAAAAATATGATGTCGGCATGAATTTCTCGACATTAACACAT 2559

Db 9177 GACAGCTATGCTTATATGAAAAAATATGATGTCGGCATGAATTTCTCGACATTAACACAT 9236

Qy 2560 GATTGGATCGAGAAATCAATGCGCATCCACCATTTTAAAAAGCTGATTTAAAAACCTTATTTT 2619

Db 9237 GATTGGATCGAGAAATCAATGCGCATCCACCATTTTAAAAAGCTGATTTAAAAACCTTATTTT 9296

Qy 2620 AATGCAATGACTTAAAGAGTATGAATGTGAAAGGGGCATCACAGGTATGTTTATGAAG 2679

Db 9297 AATGCAATGACTTAAAGAGTATGAATGTGAAAGGGGCATCACAGGTATGTTTATGAAG 9356

Qy 2680 TATGCGCTACCGCATGAGCTTCTGACGATTATTAAAGAGTTCATCACATCTCGCCAATCA 2739

Db 9357 TATGCGCTAGCGCATGAGCTTCTGACGATTATTAAAGAAAGTTCATCACATCTCGCCAATCA 9416

Qy 2740 ATTGATAGTGTGCCAGAAATATAACACTGAGGATATTTGGTTCCAAATTTGCACCTTTTAATC 2799

Db 9417 ATTGATAGTGTGCCAGAAATATAACACTGAGGATATTTGGTTCCAAATTTGCACCTTTTAATC 9476

Qy 2800 TTAGAAAAGAAAACCGGCCATGTATTATTAATAAACATCGACCTGACTTATATGCGCTTGG 2859

Db 9477 TTAGAAAAGAAAACCGGCCATGTATTATTAATAAACATCGACCTGACTTATATGCGCTTGG 9536

Qy 2860 GAACCAAAATTTACAATGAGCAAAATGAACAAATTTCAAAGTGCAGAAAACGCGAAAATATC 2919

Db 9537 GAACCAAAATTTACAATGAGCAAAATGAACAAATTTGAAAGTGCAGAAAAGAGAGAAAATATA 9596

Qy 2920 CCGGTTAAACAGTTTCATTATTAAATAGTAACGCTATAAACATTTTGCATTTTATTTAAAA 2979

Db 9597 CCGGTTAAACAGTTTCATTATTAAATAGTAACGCTATAAACATTTTGCATTTTATTTAAAA 9656

RESULT 5

AF036004

LOCUS

DEFINITION

AF036004

AF036004

AF036004.1

GI:3043922

cds.

Pasteurella multocida.

Pasteurella multocida.

Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

Pasteurella.

1 (bases 1 to 2937)

DeAngelis,P.L., Jing,W., Drake,R.R. and Achyuthan,A.M.

Identification and molecular cloning of a unique hyaluronan

synthase from Pasteurella multocida

J. Biol. Chem. 273 (14), 8454-8458 (1998)

98192845

PUBMED

9525958

2 (bases 1 to 2937)

DeAngelis,P.L., Jing,W. and Achyuthan,A.M.

Direct Submission

Submitted (26-NOV-1997) Biochem. & Molec. Biol., Univ. of Oklahoma

Health Sciences Center, 940 Stanton L. Young Blvd., Oklahoma City,

OK 73104, USA

Location/Qualifiers

1. .2937

/organism="Pasteurella multocida"

/strain="P-1059; ATCC 15742"

/db_xref="ATCC:15742"

/db_xref="taxon:747"

/note="Carter Type A strain"

19. .2937

/gene="PmHAS"

/note="gene identified by transposon insertional

mutagenesis and verified by functional expression in

heterologous cells"

19. .2937

/gene="PmHAS"

/function="polymerizes hyaluronan (HA, hyaluronate,

hyaluronic acid) polysaccharide using UDP-GlcA and

UDP-GlcNAc precursors"

/note="glycosyltransferase; HA synthase (synthetase);

membrane-bound enzyme"

/codon_start=1

/transl_table=11

/product="hyaluronan synthase"

/protein_id="AAC38318.1"

/db_xref="GI:3043923"

/translation="MNTLSQAIKAYNSNDYQALKLFKSAEYGRKIVEFQITKQEE

KLSAHPSVNSAHLVSNKEEVKVCDSPLDIATQLLLSNVKKVLSDSEKNTLKNKKWL

LTEKSENAEYRAVALYKDPKDLVLAPLDHVNDFWYKKRKLRLGKPEHOHVGL
SIIYTFNRPAILSLITACLVNQTHPIFFEVIVDDGSOEDLSPLIBOVENKLDIRVY
KOKNGFOASARNMGLAKLDFIGLLDCDMAFNPLWVHSYVLAELDEDDLTIGPR
KYIDTHIDPKDFLNNASLESPLFVNTNNSAAKGEVTSVLDWRLEQFEKTEINLRLS
DSRPFFAAGNVAFAKWLKNSGFFDEEFNWGGDEVEFGRFLRYGFFFTIDIGIMA
YHBPFGKEDTDEAGNIYLDIMREKVPYIRKLLPDIEDSHINRVPLSYIYPAYN
CANYLORCVDSALNOTVVDLEVCICNDGSDNTLEVINKLGNPNPRVIRMSKPNPGGIA
SASNAVSAFAGYIGOLDSDYLEPDVAELCLKEFLDKDLACVYITNKNVNDPDSL
IANGYNWPEFSREKLTAMIAHHFMFTIRAWHLTDGNEKLENAVDYDMELKLESEVG
KFHLNKICYARLVHGDNTSKKLGIOKQNHVNVYNSLRQGGITYYNDDEDDLDES
RKYIFNKAYEQEEDILDKDIKIQNKDAKIAVSIFYPNTLNGVLKLNINIEYNKNI
FVILVHVKNHILTPDKKEILAFYKHQVNIILLNNDISYTSNRILKEAHLNINKLI
SOLNLCVSIIPDNHDSLPVKNDSVAYMKYDVGMNFSAITHDWLEKTAHPFPKKLI
KTYFNNDLKSNNVKGASOGEMFTVALAHELLTIIKEVITSQSDSPEXNTDIWF
QFALLILEKKYGHVFNKSTLYMPWERKLOWTNEQIESAKRGENIPVNFKFIINSITL

BASE COUNT 1077 a 482 c 509 g 869 t
ORIGIN

Query Match 76.6%; Score 2283.4; DB 1; Length 2937;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 2540; Conservative 0; Mismatches 376; Indels 21; Gaps 1;

QY 43 ATTTTAAAGGAAGAAATGATACATTATCAAGCAATAAAGCATATAACGCAAT 102
DB 1 ATTTTAAAGGACAGAAATGATACATTATCAAGCAATAAAGCATATAACGCAAT 60
QY 103 GACTATGAATAGACTCAAAATATTGGAGAGTCTGCTGAACCTACGGCGGCAAAATC 162
DB 61 GACTATCAATTAGACTCAAAATATTGGAGAGTCTGCTGAACCTACGGCGGCAAAAT 120
QY 163 GTTGAATTCCAATATCAATGTAAGAAAACTCTCGACCAATTC- - - - - 209
DB 121 GTTGAATTCCAATATCAATGTAAGAAAACTCTCGACCAATTC- - - - - 180
QY 210 - - - - -TTATGTAAGTGAAGATAAAGAAAAACAGTGTGTCGATAGCTCATATGATATC 261
DB 181 GCACATCTTCTGTAATAAAGAAAAAGTCAATGTTTCGATAGCTCGGTAGATATT 240
QY 262 GCAACAGCTCTTACCTTCCAGGTAAAGAAATTAATCTATCCGAATCAGAAAAAATC 321
DB 241 GCAACAGCTCTTACCTTCCAGGTAAAGAAATTAATCTATCCGAATCAGAAAAAATC 300
QY 322 AGTTTAAAAATAATGAAATCTATCACTGGGAAAAATCGGAGAACGAGAAATCAGA 381
DB 301 ACGTTAAAAATAATGAAATCTCTCACAGAGAAATCTGAAATCGGAGGTAGA 360
QY 382 AAGGTGAACCTAGTACCAGAAATTTTCCTAAAGATCTGTTCTGTCCTCATTCGCCAGAT 441
DB 361 CGCGTCGCCCTTGTACCAAAAGATTTTCCCAAGATCTGTTTACGCCCTTTACCTGAT 420
QY 442 CATGTTAATGATTTTACATGTCACAAATCGAAAAAAGCTTAGGTATAAAGCCTGTA 501
DB 421 CATGTTAATGATTTTACATGTCACAAAGCAAGCAAAAGAAAGACTTGGCAATAAAGCTGAA 480
QY 502 AATAAGATATCGGTCTTCTTATTTATTTATTCCTACATTAATTCGTAGCCGTATTTAGAT 561
DB 481 CATCAACATGTTGCTCTTCTTATTCGTACACATTTCAATCGACCAAGCAATTTATCG 540
QY 562 ATAAGTTAGCTGTTGGTCAATCAGAAAAAACAATACCATTTGAGAGTCTGTTGCA 621
DB 541 ATTTACATTTAGCTGTTTGTAGTAAACCAAAACACATTTACCCGTTTGAAGTTATCGTGACA 600
QY 622 GATGATGCTAGTGAAGAAAACTTACTTACCATTGTCGAAAAATACGAAAAAATCTTGAC 681
DB 601 GATGATGCTAGTGAAGAAAACTTATCACCAGTATCTCCCAATATGAAAAATTTGGAT 660
QY 682 ATAAAAGTATGAAGAAAAAGATTTATGGATATCAATTTGTGAGAGTCAAGAACTTAGT 741
DB 661 ATTCGCTACGTACAGAAAAAGATAACGGTTTTCAGCCAGCTGCGGCTCGGAATATGGA 720
QY 742 TTACGTACAGAAATGATGATTTCTCTCGATTCCTAGACTGCGATATGCAACCAACAA 801
DB 1801 ATTAGAGCTTGCGATTTTAACTGATGATCAATGAAAAAATTTGAAATGCCGTAGACTAT 1860

DB 721 TTACGCTTAGCAAAATATGACATTTATTTGGCTTACTCGACTGTGATATGGCGCAAAATCCA 780
QY 802 TTATGGGTTTCATTTCTTATCTTACAGAACTATTAGAAGCAATGATATTTTAAATTTGA 861
DB 781 TTATGGGTTTCATTTCTTATGTTGCAGAGCTATTAGAAGATGATGATTAAACAATCATTTGT 840
QY 862 CCTAGAAAAATATGTTGCTACTCATTAATATTACCGGAGAGCAATTTCTTAAACGATCCATAT 921
DB 841 CCAAGAAAAATACATCGATACACAACTATTGACCCAAAAGACTTCTTAAATAACGGGAGT 900
QY 922 TTAAATAGATCACTACTCTGAAACCGCTACAAATTAACAATCTTTCGATTTACATCAAAAGGA 981
DB 901 TTGCTTGAATCATTTACAGCAAGTGAACCAATTAATAGTGTCCCGCAAAAGGGAAGGA 960
QY 982 AATATATCGTTGGATTGGAGATTAGAACATTTCAAAAAACCGATATCTACGCTCTATGT 1041
DB 961 ACAGTTTCTCTGGATTGGCGCTTAGAACAAATTCGAAAAACAGAAAAATCTCCGCTTATCC 1020
QY 1042 GATTTCCGTTTCGTTTATTTTGGTGGGGTAAATGTTTGCATTTTCTAAAGAAATGGCTAAAT 1101
DB 1021 GATTCGCTTTCCGTTTATTTTGGCGGGTAAATGTTTGCCTTAAAGAAATGGCTAAAT 1080
QY 1102 AAAGTAGGTTGTTTCGATGAAGAAATTTAAATCATTTGGGGGGCGGAAGATGTAGAAATTTGGT 1161
DB 1081 AAATCCGGTTTCTTGTAGAGAGAAATTTAATCACTGCGGTGGAGAGATGTGGAATTTGGA 1140
QY 1162 TACAGATTATTGCAAAAGGCTGTTTTTTCAGAGTAAATGACGGCGGAATGGCCATCAT 1221
DB 1141 TATCGCTTATTCGTTTACGTTAGTCTTTTAAACTATTGATGGCATTTATGGCTTACCAT 1200
QY 1222 CAAGAACCCCTGGTAAAGAAATGAACAGACGCGAAGCTGCTAAAGTATTATACGCTT 1281
DB 1201 CAAGAGCCAGCAGGTAAAGAAATGAACCGGATCGTGAAGCGGGAAGAAATATTACGGCTC 1260
QY 1282 AAAATCTGAAAGAAAAAGTACCTTACATCTATAGAAAGCTTTTACCATAAGAGATTCA 1341
DB 1261 GATATTATGAGAGAAAAAGTCCCTTATATCTATAGAAACTTTTACCATAAGAGATTCG 1320
QY 1342 CATATTATAGAAATACCTTTAGTTTCTTATTTATATCCCGCTTATACTGTGCAAAATTTAT 1401
DB 1321 CATATAATAGAGTACCTTTAGTTTCTTATTTATATCCAGCTTTAATCTGTGCAAACTAT 1380
QY 1402 ATTCAGAGATGTTAGATAGTGTCTTAACTCAAACTGTTGCTGATCTCGAGGTTTGTATT 1461
DB 1381 ATTCAGAGTTCGGTAGATAGTGTGCTGATGACTGATGACTGTTGTTGATCTCGAGGTTTGTATT 1440
QY 1462 TGTAAACGATGTTTCAACAGATTAATACCTTAGAAGTGAATCAATAAGCTTTATGTTATTAAT 1521
DB 1441 TGTAAACGATGTTTCAACAGATTAATACCTTAGAAGTGAATCAATAAGCTTTATGTTATTAAT 1500
QY 1522 CCTAGGTTACGATCATGCTTAAACCAATGGCGGAATAGCTCAGCATCAAAATCGAGCC 1581
DB 1501 CCTAGGTTACGATCATGCTTAAACCAATGGCGGAATAGCTCAGCATCAAAATCGAGCC 1560
QY 1582 GTTCTTTTGTAAAGGTTTATTTACATTTGGGAGTTAGATTTCAGATGATTATCTTGAGCCT 1641
DB 1561 GTTCTTTTGTAAAGGTTTATTTACATTTGGGAGTTAGATTTCAGATGATTATCTTGAGCCT 1620
QY 1642 GATGAGTTTGAACCTGTTTAAAAGAAATTTTAAAAGATAAAGCTAGCTTGTGTTTAT 1701
DB 1621 GATGAGTTTGAACCTGTTTAAAAGAAATTTTAAAAGATAAAGCTAGCTTGTGTTTAT 1680
QY 1702 ACCACTAATAGAAACGTCATCCCGGATGTTAGCTTAAATGCTTAATGGTTTACAAATTTGGCCA 1761
DB 1681 ACCACTAATAGAAACGTCATCCCGGATGTTAGCTTAAATGCTTAATGGTTTACAAATTTGGCCA 1740
QY 1762 GAATTTTCCAGGAAAAAATCTACAAAGGCTATGATTGCTTACCAATTTTGAATGTTTATAG 1821
DB 1741 GAATTTTCCAGGAAAAAATCTACAAAGGCTATGATTGCTTACCAATTTTGAATGTTTATAG 1800
QY 1822 ATTAGAGCTTGCGATTTTAACTGATGATTTAACGGAATTTAACGGAATTTTAAACCCGCTGGAAT 1881
DB 1801 ATTAGAGCTTGCGATTTTAACTGATGATTTAACGGAATTTTAAACCCGCTGGAATTTTAACT 1860

Qy	1882	GACATGTTCTTAACTCAGTGAAGTTGGAAATTTTAAACATCTTATAAAATCTCGTAT	1941
Db	1861	GACATGTTCTTAACTCAGTGAAGTTGGAAATTTTAAACATCTTATAAAATCTCGTAT	1920
Qy	1942	AACCGGTATTTACATGTCATTAACATCCATTAAAGAAATCGCATTTCAAAAGAAAAC	2001
Db	1921	AACCGGTATTTACATGTCATTAACATCCATTAAAGAAATCGCATTTCAAAAGAAAAC	1980
Qy	2002	CATTTTGTGTAGTCAATCAGTCATTAAATAGACAAGGCATCAATATTATAAATTATGAC	2061
Db	1981	CATTTTGTGTAGTCAATCAGTCATTAAATAGACAAGGCATCAATATTATAAATTATGAC	2040
Qy	2062	AAATTTGATGATTTAGATGAAAGTAGAAGATATATCTTCAATAAAACCGTGAATATCAA	2121
Db	2041	GAATTTGATGATTTAGATGAAAGTAGAAGATATATCTTCAATAAAACCGTGAATATCAA	2100
Qy	2122	GAAGAAATGGATGTTTAAAGATCTTAAACATCTTAAACATTAAGATGGCCAAATCGCA	2181
Db	2101	GAAGAAATGGATGTTTAAAGATATTTAAATCATCCAGAATTAAGATGGCCAAATCGCA	2160
Qy	2182	GTGAGTATTTTCTATCCCAATACATTAACCGGCTTAGTGAAAAAATCAACAAATATTATT	2241
Db	2161	GTGAGTATTTTCTATCCCAATACATTAACCGGCTTAGTGAAAAAATCAACAAATATTATT	2220
Qy	2242	GAATATAATAAAATATATTCGTTATTTCTACATGTTGATAGAATCATCTTACACCA	2301
Db	2221	GAATATAATAAAATATATTCGTTATTTCTACATGTTGATAGAATCATCTTACACCA	2280
Qy	2302	GACATCAAAAAGAAATATTGGCTTTCTATCATAGCACCAAGTGAATATTTTACTTAAT	2361
Db	2281	GATATCAAAAAGAAATATTGGCTTTCTATCATAAACATCAAGTGAATATTTTACTTAAT	2340
Qy	2362	AATGACATCTCATATTACACGAGTAATAGACATAATFAAAACTGAGGCACATTTAAGTAAT	2421
Db	2341	AATGATATCTCATATTACACGAGTAATAGATTAATAAACTGAGGCATTTAAGTAAT	2400
Qy	2422	ATTAATAAAATTAAGTCAGTTAAATCTAAATTGTGAATACATCATTTTGTGATAATCATGAC	2481
Db	2401	ATTAATAAAATTAAGTCAGTTAAATCTAAATTGTGAATACATCATTTTGTGATAATCATGAC	2460
Qy	2482	AGCCTATTCGTTAAAAATCACAGCTATGCTTTATATGAAAAATATGATGTCGGCATGAAT	2541
Db	2461	AGCCTATTCGTTAAAAATCACAGCTATGCTTTATATGAAAAATATGATGTCGGCATGAAT	2520
Qy	2542	TTCTCAGCATTAACACATGATTTGGATCGAGAAAATCAATGGCATCCACCATTTTAAAAAG	2601
Db	2521	TTCTCAGCATTAACACATGATTTGGATCGAGAAAATCAATGGCATCCACCATTTTAAAAAG	2580
Qy	2602	CTGATTTAAACCTTTTAAATGACAAATGACTTTAAGAAGTATGAATGTGAAAGGGCATCA	2661
Db	2581	CTCATTTAAACCTTTTAAATGACAAATGACTTTAAGAAGTATGAATGTGAAAGGGCATCA	2640
Qy	2662	CAAGGTATGTTTATGAAGTATCGCTACCGCATGAGCTTCTGACGATTTATAAGAAGTCT	2721
Db	2641	CAAGGTATGTTTATGAAGTATCGCTACCGCATGAGCTTCTGACGATTTATAAGAAGTCT	2700
Qy	2722	ATCACATCTCGCCAATCAATTTGATAGTGGCCAGAAATATAACACTGAGGATATTGGTTC	2781
Db	2701	ATCACATCTCGCCAGTCAATTTGATAGTGGCCAGAAATATAACACTGAGGATATTGGTTC	2760
Qy	2782	CAATTTGCATCTTTTAAATCTTAGAAAAGAAAACCGGCCATGTATTTAATAAACATCGACC	2841
Db	2761	CAATTTGCATCTTTTAAATCTTAGAAAAGAAAACCGGCCATGTATTTAATAAACATCGACC	2820
Qy	2842	CTGACTTATATGCTTTGGGAACGAAAATTTACAATGGACAAATAGAACAAATTTCAAAGTGCA	2901
Db	2821	CTGACTTATATGCTTTGGGAACGAAAATTTACAATGGACAAATAGAACAAATTTCAAAGTGCA	2880
Qy	2902	AAAAAGGCGAAAATATCCCGTTAACAGTTCATTATTATAGTATAACGCTATAA	2958
Db	2881	AAAAAGGAGAAAATATCCCTGTTAACAGTTCATTATTATAGTATAACCTATAA	2937

```

RESULT 6
AF237926      2919 bp      DNA      linear      BCT 08-MAY-2000
LOCUS
DEFINITION   Pasteurella multocida hyaluronan synthase (has) gene, complete cds.
ACCESSION   AF237926
VERSION      AF237926.1 GI:7716512
KEYWORDS
SOURCE
ORGANISM
Pasteurella multocida.
Pasteurella multocida
Pasteurella
Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
1 (bases 1 to 2919)
Fuller,T.E., Kennedy,M.J. and Lowery,D.E.
Identification of Pasteurella multocida virulence genes in a
septicemic mouse model using signature-tagged mutagenesis
Unpublished
2 (bases 1 to 2919)
Fuller,T.E., Kennedy,M.J. and Lowery,D.E.
Direct Submission
Submitted (24-FEB-2000) Discovery Research, Pharmacia & Upjohn
Animal Health, 7923-25-434, 7000 Portage Road, Kalamazoo, MI
49001-0199, USA
FEATURES
source      1..2919
              Location/Qualifiers
              /organism="Pasteurella multocida"
              /db_xref="taxon:747"
              1..2919
              /gene="has"
              1..2919
              /gene="has"
              /note="pmhas"
              /codon_start=1
              /transl_table=11
              /product="hyaluronan synthase"
              /protein_id="AAF68412.1"
              /db_xref="GI:7716513"
              /translation="MNTLSQAIKAYNSNDYQLAKLFEKSAEYGRKIVVEFOITKQCE
              KLSAHVSNAHLSVKNKEKRVNCDSPDIAOTLLSNVKKLVLSSEKNTLKKKKWL
              LTBKSENAEVAVALVDFKDPKDLVLPDPHNDFTWYKKRKRIGIKPEHGHVRY
              SIIVTFNPAILSIITLACLVNOKTHYPEVIVTDGSGEDLSPIIRQYENKLDIRYV
              RQINDQFASACRMGLRAKYDFIGLLCDMAPNPLWVHSYVAELEDDDLITIGPR
              KIDTOHIDPKDFLNNAISLDSLEUPEVTNNVSAAGKEGTVSLDMRLQEQFETENLRS
              DSPERFQANVAFKWLKNSGFFDEEFGNHMGGEVDFGRFLRYGSRFKITDGLMA
              YHQEPGKETDREAGKNTLDIRKPYIYKLLPIEDSHINRVLVSIYIPAN
              CANTQRCVDSALNQTVVLEVCINDGSDTNTLEFKLKDKTACVYTTNRNRPNGGIA
              SASNAVFAFGYIIGOLDSDYDPAVELCVLEKLTACVYTTNRNRPNGDGLS
              IYANGNWPESRKLTKAMTAHFHRMFTIRAWHLTDGFEKIEINADYDMFLKLSVG
              KFKHLNICYVNRVHGDNTSIKIGLKQKNHVVVQNSNRQITVYNTDFDLDLSEV
              RYVIFNKATQEEIDILKIKIIONKDAIKVIFPNTLGLVKLLNNIIEYNNKI
              FVLVLHVDKWHLPDIIKKEILAFYHKHVNLIANNDSIVYTSNRLIKTBAHLSNINKL
              SOLNACEYILFNDHDSLFVKVNSYAYMKKYDVGMPFSALTHDWIEKINAHPPFKKLI
              KTFYNDNDLAKSMHVGQGMFTYLAHELLTIIIRKIVTSCQSDISQVPEINTDFIWF
              QFALLILEKKTGHTVNTKSTLTLYTPMWERKLQWTEQIESAKRGENIPVNRKFIINSIT
BASE COUNT      1059 a      481 c      506 g      863 t
ORIGIN
Query Match      76.2%; Score 2268.6; DB 1; Length 2919;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 2524; Conservative 0; Mismatches 374; Indels 21; Gaps 1;
Qy 61 ATGAATACATTATCACAGCAATAAAGCATATATACAGCAATGACTATGAATTAGCACTC 120
      |||||
Db 1 ATGAATACATTATCACAGCAATAAAGCATATATACAGCAATGACTATGAATTAGCACTC 60
      |||||
Qy 121 AAATATTGAGAACTGCTGAACACCTACGGCGGAGAAAATTCGTTGAATTCCAAATTCATC 180
      |||||
Db 61 AAATATTGAGAACTGCGCGGAATCTATGACGCGGAAAATTCGTTGAATTCCAAATTCAC 120
      |||||
Qy 181 AAATGTAAGAAAAACTCTCGACCAATTCCTTATG-----TAAAT 219
      |||||
Db 121 AAATGCCAAGAAAAACTCTGACGACCTCTCTTCTTAATTCAGCACATCTTTCTGTGAAT 180

```

Qy 220 GAAGATAAAAAACAGTGTTTGGATAGCTCATTAGATATCGCAACACAGCTCTTACTT 279
Dy 181 AAAGAGAAAGTCAATGTTTGGATAGTCGGTTAGATATGCAACACACTGTTACTT 240
Qy 280 TCCACGCTAAATAATTAACCTCTTCCGATCAGAAAAACAGTTTAAAAATAAATGG 339
Dy 241 TCCACGCTAAATAATTAAGTCTTCTGACTCGGAAAAAACACGTTTAAAAATAAATGG 300
Qy 340 AAATCTACACTGGGAAAAAATCGGAGAACGACAGAAATCAGAAAGGTGGACTAGTACCC 399
Dy 301 AAATGCTCACTGAGAGAAATCTGAAATCGGAGGTGAAGCGGTGCGCCCTGTATCCA 360
Qy 400 AAAGATTTCTTAAGATCTTGTCTCTGCTCCATTGCCAGATCATGTTAAATGATTTTACA 459
Dy 361 AAAGATTTCCCAAGATCTGTTTACGGCTTTACCTGATCATGTTAAATGATTTTACA 420
Qy 460 TGGTACAAAAATCGAAAAAAGCTTAGGTATAAAGCCTGTAATTAAGAAATACGGTCTT 519
Dy 421 TGGTACAAAAAGCGAAAAAGACTTGGCATAAAACTGAACATCAACATGTTGGTCTT 480
Qy 520 TCTATTATATTCCTACATTTAATCGTAGCGGTATTTTAGATATTAACGTTAGCCGTTTG 579
Dy 481 TCTATTATCGTTACACATTTCAATCGACGACCAATTTTATCGATTTACATTTAGCCCTGTTTA 540
Qy 580 GTCAATCAGAAAAAATACCTACCTATTCGAAGTCGTTGTGCAGATGATGGTAGTAAGGAA 639
Dy 541 GTAACCAAAAAACATATACCGCTTGAAGTTATCGTGACAGATGATGGTAGTAGGAA 600
Qy 640 AACTTACTTACCATGTGCAAAAAATACGAAACAAAAACCTTGACATAAAGATATGTAAGACAA 699
Dy 601 GATCTATCACCGATCATTCGGCAATATGAAATAAATTTGGATATTCGCTAGCTCAGACAA 660
Qy 700 AAAGATTTAGATATCAATGTGTGCTCAGTACAGAACTTAGTTTACGTACAGCAAGTAT 759
Dy 661 AAAGATAACGGTTTCAAGCCAGTCCCGCTCGGAATATGGGATTAGCTTAGCAAAATAT 720
Qy 760 GATTTTGTCTCGATTTCTAGCTGCGATATGGCACACACAAACATTTATGGTTTCATCTTAT 819
Dy 721 GACTTTATTTGGCTTACTCGCTGCTGATATGGCGCAAAATCCATTTATGGTTTCATCTTAT 780
Qy 820 CTTACAGAACTATTAGACACATGATATTTTAAATGGACCTAGAAAAATATGTGGAT 879
Dy 781 GTTGCAAGACTATTAGAAGATGATGATTAACAATCATTTGTGTCGAAGAAATACATCAT 840
Qy 880 ACTCATATATTACCGCAGACAAATTCCTTAACGATCCATATTTAATAGAATCACTACCT 939
Dy 841 ACACAACTATTGACCCAAAGACTTCTTAATAACGCGAGTTTCTGTTGAATCATATCCA 900
Qy 940 GAAACCGCTACAAATAACAACTCTTCGATTTACATCAAAAGGAAATATATCGTTGGATTGG 999
Dy 901 GAAGTGAAACCAATAATAGTGTCCCGCAAAAGGGAGAACAGTTTCTCTGGATTGG 960
Qy 1000 AGATTAGAACATTTCAAAAAACCGATTAATCTACGTCTATGATGATTCGCGTTTCGTTAT 1059
Dy 961 CGCTTAGAACATTTGAAAAACAGAAAAATCTCCGCTTATCGGATTCGCGTTTTCGTTTT 1020
Qy 1060 TTTGTTGCGGTAATGTTGCTATTTTCAAGAAATGCGTAAATAAAGTAGTTGTTTCGAT 1119
Dy 1021 TTTGCGCGGGTAATGTTGCTTTCCTCAAAATGGCTAAATAAATTCGTTTCTTTGAT 1080
Qy 1120 GAAGATTTAATCATTTGGGGGGCGGAAGATGTAGAATTTGGTTTACAGATATTTTCCCAA 1179
Dy 1081 GAGGAATTTAATCACTGGGGTGGAGAAGATGTGGAATTTGGATATTCGCTTATTCGGTTAC 1140
Qy 1180 GCGTGTTTTTCAGAGTAAATTCACGGCGGAATGGCCATCCATCAAGAACCCCTGGTAA 1239
Dy 1141 GGTAGTTTCTTTAAAACATTTGATGGCAATTTATGGCTACCATCAAGAGCCACAGGTAAA 1200
Qy 1240 GAAATGAACAGAACGCGAAGCTGGTAAAGTATTTAGCTTTAAATTTGTGAAAGAAAG 1299
Dy 1201 GAAATGAACCCGATCGTGAACGGGAAAAAATATTTACGCTCGATTTATGAGAGAAAG 1260

Qy 1300 GTACCTTACATCTATAGAAAGCTTTTACCAATAGAAAGATTACATATTTCAATAGATACCT 1359
Dy 1261 GTCCCTTATATCTATAGAAAACTTTTACCAATAGAAAGATTGCATATCAATAGAGTACCT 1320
Qy 1360 TTAGTTTCTATTTATATCCCGCTTATTAACCTGCAAAATTAATTTCAAGATGTGTAGAT 1419
Dy 1321 TTAGTTTCTAAATTTATATCCAGCTTATAACTGTGCAAACTATATTCAAGTTGCGGTAGAT 1380
Qy 1420 AGTGTCTTTAATCAAACTGTTTGTGATCTCGAGGTGTTGTTTGTGTAACGATGTTTCAACA 1479
Dy 1381 AGTGCACTGAATCAGACCTGTTGTTGATCTCGAGGTTGTTGTTGTAACGATGTTTCAACA 1440
Qy 1480 GATAATACCTTAGAAGTGAATCAATAGCTTTATGTTAATTAATCTTAGGGTACGATCATG 1539
Dy 1441 GATAATACCTTAGAAGTGAATCAATAGCTTTATGTTAATTAATCTTAGGGTACGATCATG 1500
Qy 1540 TCTAAACCAATGGCGGAATAGCCTCAGATCAAAATGACGCGTTTCTTTTCTTAAAGT 1599
Dy 1501 TCTAAACCAATGGCGGAATAGCCTCAGATCAAAATGACGCGTTTCTTTTCTTAAAGT 1560
Qy 1600 TATTACATTTGGCAGTTAGATTCAGATGATTTTGTGAGCCTGATGCAAGTTTCAACTGTCT 1659
Dy 1561 TATTACATTTGGCAGTTAGATTCAGATGATTTTGTGAGCCTGATGCAAGTTTCAACTGTCT 1620
Qy 1660 TTAAAGAAATTTTAAAAAGATAAAACGCTAGCTTGTGTTTATACCACTAATAGAAACGTC 1719
Dy 1621 TTAAAGAAATTTTAAAAAGATAAAACGCTAGCTTGTGTTTATACCACTAATAGAAACGTC 1680
Qy 1720 AATCCGGATGGTAGCTTAAATCGCTAATGGTTACAAATTTGCCAGAAATTTTACAGAGAAAA 1779
Dy 1681 AATCCGGATGGTAGCTTAAATCGCTAATGGTTACAAATTTGCCAGAAATTTTACAGAGAAAA 1740
Qy 1780 CTCACACCGCTTATGATTTGCTCACCATTTTAGAAATGTTTACGATTTAGAGCTTTGGCATTTA 1839
Dy 1741 CTCACACCGCTTATGATTTGCTCACCATTTTAGAAATGTTTACGATTTAGAGCTTTGGCATTTA 1800
Qy 1840 ACGGATGGATTTAAGCAAAATTAATGAAACGCGGTGGATTTAGACATTTTCTTAAACTC 1899
Dy 1801 ACTGATGGATTTCAATGAAAAATTTGAAATGCGGTAGACTATGACATGTTCTTCAAACTC 1860
Qy 1900 AGTGAAGTTGGAAAAATTTAAACATCTTAATAAAATCTGCTATAACGCGCTATTACATGCT 1959
Dy 1861 AGTGAAGTTGGAAAAATTTAAACATCTTAATAAAATCTGCTATAACGCGGTATTACATGCT 1920
Qy 1960 GATAACACATCCATTAAGAAACTCGGCATTTCAAAAGAAAAACCATTTTGTGTAGTCAAT 2019
Dy 1921 GATAACACATCAATTAAGAAACTTGSCATTTCAAAAGAAAAACCATTTTGTGTAGTCAAT 1980
Qy 2020 CAGTCATTAATAGACAGGCATCAATTAATTAATTAATGACAAATTTGATGATTTAGAT 2079
Dy 1981 CAGTCATTAATAGACAGGCATCAATTTAATTAATTAATGACAAATTTGATGATTTAGAT 2040
Qy 2080 GAAAGTAGAAAGTATCTTCAATAAAACCGCTGAATATCAAGAGAAATGATATGTTA 2139
Dy 2041 GAAAGTAGAAAGTATCTTCAATAAAACCGCTGAATATCAAGAGAGATTTGATATCTT 2100
Qy 2140 AAAGATCTTAAACTCATTTCAAAATAAAGATGCAAAATCGCAATCGAGTCAGTATTTTCTATCCC 2199
Dy 2101 AAAGATTTAATCATCCAGATTAAGATGCCAAATCGCAATCGAGTCAGTATTTTATCCC 2160
Qy 2200 AATACATTAACCGCTTAGTGAATAAACTAAACATATTTATGAAATATTAATAAATAATA 2259
Dy 2161 AATACATTAACCGCTTAGTGAATAAACTAAACATATTTATGAAATATTAATAAATAATA 2220
Qy 2260 TTTGCTTATTTCTACATGTTGATAGAAATCATTTTACACACAGACATCAAAAGAAATA 2319
Dy 2221 TTTGCTTATTTCTACATGTTGATAGAAATCATTTTACACAGATATCAAAAGAAATA 2280
Qy 2320 TTTGCTTCTTATCATAAGCACCAGTGAATTTTACTAAATAATGACATCTCATATTTAC 2379
Dy 2281 CTAGCCTTCTCATATAAACATCAAGTGAATTTTACTAAATAATCATATCTCATATTTAC 2340
Qy 2380 ACGAGTAATAGACTAATAAAACTGAGGCACATTTTAAGTAAATTAATTAATAAGTCAG 2439

Db	8973	TGCAATGATATGTCATCATTTTCAGGATGTTCCACAGCAAGACGATGGAAACCTTAACCTGAAGG	8914		
Qy	1848	ATTTTAACGCAAAATATTTGAAAACGCGGTGGATTATGACATGTTCCTTTAAACCTCAGTGAAGT	1907		
Db	8913	TTTTCAACGATCGATCAGCAACGCGAGTTGATTACGATGATGTATTTTAAACCTTAGTGAAGT	8854		
Qy	1908	TGGAATAATTTAAACATCTTTAATAAATCTGCTATTAACCGCGTATTACATGCGTGATAACAC	1967		
Db	8853	TGGAACCGTTCAAGCATATAAACAATAATTTGTATATATCGCGTATTTGCAATGGTGAATAAC	8794		
Qy	1968	ATCCATTAAAGAACTCGCGATTTCAAGCAAAAACCATTTTCTGCTAGTCAATCAGTCAAT	2027		
Db	8793	GTCTATAAAAAGTTGGATATTTCAAGGAAGAAATCATTTTAAAGTTGTTAACGAATCAAT	8734		
Qy	2028	AAATAGACAAGGCATCAATTTATTAATATGACAAAATTTGATGATTTAGATGAAGTAG	2087		
Db	8733	AAGTAGGCTAGGCATATAAAAAATATAAATATTCACCACTAACTAATTTTGAATGAATGTAG	8674		
Qy	2088	AAAGTATATCTTCAATTAACCGCTGAATATCAAGAAGAAATG	2130		
Db	8673	AAATATACCTGGGAAAAAATAGAGAATGATTTATAATTTATG	8631		
RESULT 8					
AC007926/c					
LOCUS					
DEFINITION		AC007926 155204 bp DNA linear HTG 17-JUL-2001			
Typanosoma brucei chromosome II clone RPC193-3H15, *** SEQUENCING					
IN PROGRESS ***, 1 ordered piece.					
ACCESSION		AC007926			
VERSION		AC007926.8 GI:14787210			
KEYWORDS		HTG; HTGS_PHASE2.			
SOURCE		Typanosoma brucei.			
ORGANISM		Typanosoma brucei			
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;					
Typanosoma.					
REFERENCE					
AUTHORS					
El-Sayed,N.M., Ghedin,E., Song,J., Larkin,C., Wanless,D., Jones,K., Peterson,S., Hou,L., Zhao,X., Mason,T., Millscher,J., Pai,G., Van Aken,S., Utterback,T., Khalak,H.G., Gerard,C., Leech,V., Ullu,E., Melville,S., White,O., Adams,M.D., Donelson,J.E. and Fraser,C.M.					
Typanosoma brucei GUTat10.1 RPC193-3H15 BAC genomic sequence					
TITLE					
Unpublished					
REFERENCE					
AUTHORS					
El-Sayed,N.M., Khalak,H. and Adams,M.D.					
TITLE					
Direct Submission					
JOURNAL					
Submitted (28-JUN-1999) The Institute for Genomic Research, 9712					
Medical Center Dr, Rockville, MD 20850, USA					
COMMENT					
On Jul 17, 2001 this sequence version replaced gi:12746529.					
* NOTE: This is a 'working draft' sequence. It currently					
* consists of 1 contigs. Gaps between the contigs					
* are represented as runs of N. The order of the pieces					
* is believed to be correct as given, however the sizes					
* of the gaps between them are based on estimates that have					
* provided by the submittor.					
* This sequence will be replaced					
* by the finished sequence as soon as it is available and					
* the accession number will be preserved.					
* 1 155204: contig of 155204 bp in length.					
FEATURES					
source					
1. .155204					
/organism="Typanosoma brucei"					
/isolate="GUTat10.1"					
/db_xref="taxon:5691"					
/chromosome="II"					
/clone="RPC193-3H15"					
BASE COUNT 48747 a 31603 c 32329 g 42525 t					
ORIGIN					
Query Match 3.3%; Score 97.6; DB 2; Length 155204;					
Best Local Similarity 45.5%; Pred No. 6.8e-07;					
Matches 346; Conservative 0; Mismatches 414; Indels 0; Gaps 0;					
Qy	1912	AAATTTAAACATCTTAATAAATCTGCTATTAACCGCGTATTACATGCGTGATAACACATCC	1971		

the Mucosal Pathogen, Campylobacter jejuni. BIOSYNTHESIS OF STIALLYLATED GANGLIOSIDE MIMICS IN THE CORE OLIGOSACCHARIDE
J. Biol. Chem. 277 (1), 327-337 (2002)
11689567
2 (bases 1 to 11442)
Gilbert, M., Michniewicz, J., Karwaski, M.-F., Cunningham, A. and Wakarchuk, W.
Direct Submission
Submitted (15-JUL-2001) Institute for Biological Sciences, National Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A 0R6, Canada

FEATURES

Location/Qualifiers

1..11442
/organism="Campylobacter jejuni"

/strain="ATCC 43438"

/db_xref="ATCC:43438"

/db_xref="taxon:197"

/note="Penner-type strain O:10"

<1..348

/gene="waaC"

<1..348

/gene="waaC"

/codon_start=1

/transl_table=11

/product="heptosyltransferase I"

/protein_id="AAK91719.1"

/db_xref="GI:15193215"

/translation="FAKEVNLNGSDIETHIELAPFSLBELMAFTKMDLIIGNSGP
THLAPALNKASITIFGATPSYRNAFQTHINKIIDAGKIQNAKHIDKSDFCITRIEE
DIFKLKGLLNEK"

338..1225

/gene="htrB"

338..1225

/gene="htrB"

/codon_start=1

/transl_table=11

/product="lipid A biosynthesis acyltransferase"

/protein_id="AAK91720.1"

/db_xref="GI:15193216"

/translation="MKNSDRIYLSIYIILKFVTPMPCILHFLALIVARITFLNKK
HRKIIDINQICFPQYOKERDKLSKIYENFAQGDICLQNTTKKELKLVAFIN
ENLLDALAKRPIIFTTAHYGNWEILSLAYAAKYGATISVGGKILSEIMYEILSONR
TQFDIELIDKGGIKSALKRAGILITDQDCVEKESVRLKFNKEVNYOMGASL
IAQRNALIPVAYKEDGKFCIEFFVKVQDSQASLEELTYQAQSCDEMIKKRPWEY
FEHRRFASYNKEIYKDAK"

1222..2769

/note="orf3"

/codon_start=1

/transl_table=11

/product="putative two-domain glycosyltransferase"

/protein_id="AAK91721.1"

/db_xref="GI:15193217"

/translation="MNLKQISRIIIVNKAQVILLECILSLKDFDEILILNNESSDNTL
NTANEFKDFANLHIYHSAPIGFALKNLALSAYAKNDWILSDADAEVLENECKELKN
LEIQEDNIIILRKNLKGEWIACAGWPDYVMRIFNKNFRNDNLVHESVLPSNA
KKIYLNGRHLAFRDIHSLIDKMQYSSILWAKQIHKKSGLVKANLRAFVFFRNYF
LVNGLFYGVKGTISVNCALGTFYKMKLYELQKQPKTCALITITYNQKRLKLVLD
SKNGLAFLENLVADGSKEDTARLIEYQKDFPCPLKHWOEDGEGFRAAKSRNNAI
KASKSYIILIDGMILEKDFVKNHLEFAKRVKVLQGSVILNKNSEETLKNNYSL
AFNKGFKQRSIFLAKCVKFSKTLKFFKKKSQLYVKGSKTCNNSPKSDFEALGPN
EPIFGWGREDESEVAREFLNDGVFKRLKFNALUAIYHIHEENSKNMLEINHQIYLETIK
NEKTTWR"

2773..3957

/note="orf4"

/codon_start=1

/transl_table=11

/product="putative glycosyltransferase"

/protein_id="AAK91722.1"

/db_xref="GI:15193218"

/translation="MKKGVVPIYVNVKYLKCELDVSNQTYTNLEIILVNDGSTDE
NSNLAKYTLKDKRTLFDKKNGLSSARNVGIEVFSGEYKLKNKTOHINKENSLIEF
QLDGNPNYIKAYKSSQAFNEKDLTNLTYPSIDYIIFLDSNKNKLCIECVIRM
KNVDVLFWDHDCITIEDNINKNKKRTRMEIFDFKKECIITPEKAYANRALSVGSRDISG

gene

CDS

WNGVIDNFLKQIKLFINFINEDIHFGIILFASANKIYVLSQLYLCRLRANSISN
HDKKITRANSEYFKDIYETEGENAKENYLKAASRVITALKLEFTEPKDQSENALA
IKETFLPCYAKKALMIKKFKKDLNLKQLVLKPFQTKLPYDIWKEWQIKKYLLII
KI"
4013..5071
/gene="cgtA"
4013..5071
/gene="cgtA"
/codon_start=1
/transl_table=11
/product="beta-1,4-N-acetylglactosaminyltransferase"
/protein_id="AAK91723.1"
/db_xref="GI:15193219"

/translation="MLFQSYFVKIICLFIPFRKIRHKIKKTKELLKNIQDKIDSYPK
KTLIOINKYNNEDLJKLNAKIIIGGKHGKGFVNDKSKDPKSPINPWFARVKNEAITL
KASLESILPAIQORGVIGYNDCTDSEIILEFCQYPSGPIKIPYVEIOIQNPKESEN
KLISYNTVAFIPKDEWLKIDVDHYADKLYKSFYIPRKNYHVISYRDIQFIFNE
EKFYVRNKEGILKAPGDLAIQNTNLFWKELLIEDDFKWNATAKNENAKSYEIL
KVRNRIYFTTELNNYHFFPKYRKNQYKQLNWNVLDDDFIKNYKEKLNQIDFKMLEY
KTLKVKYKLTSSASDKI"
complement(5043..5936)
/gene="cgtB"
complement(5043..5936)
/gene="cgtB"
/codon_start=1
/transl_table=11
/product="beta-1,3-galactosyltransferase"
/protein_id="AAK91724.1"
/db_xref="GI:15193220"

gene

CDS

/translation="WFKISIIILTYNVEQYIARATESCINOTFKNTEIIVVDGSDK
SIDIVYAKKDDRIKIIHNEENKLLRARIYGVKANSFYINFLPDDYLELNACEE
CMKILKNEIDILFFNAFVLENNKNIERKLNQKCYVRKDFLKLKLNLPWTWVA
KVTKELYLKAVGLISLENAKINMAEDVLLYPLINISNTIFHLSKLNLYNQINFSI
TKTLTLQNTKNTIOQDNVLYLLKKMQYNFNLTLLKLIYEFLLEIKYSLSSKRNVL
CFKINIFPKKIQKFYRLKLM"
6025..6900
/gene="cstII"
6025..6900
/gene="cstII"
/codon_start=1
/transl_table=11
/product="alpha-2,3-/alpha-2,8-sialyltransferase"
/protein_id="AAK91725.1"
/db_xref="GI:15193221"

gene

CDS

/translation="MKKVIIISGNPSLKEIDYSRLENDFVFCNFOYFEDKYVLGKK
FRAYFYNGLFFEQYTLKHIQONQYETELIMCSNYNQAHLENENFVTFYDIFPDA
HLGYDFKQLKEFNAYFHEIYNQRTISGVYMCVAIYALGYEILSGIDFYQNGS
SYAFDTKQENLLKLAPDKNDRSHYIGHSKNTDIKALEFLEKTYKIKLYCLPNSLIA
NFIELAPLNSNFIQIEKNNTYTKDILIPSEAYGKFSKNINFKKIKIKENIYKLIKND
LLRLPSDIKHYFKGK"
6901..7941
/gene="neuB1"
6901..7941
/gene="neuB1"
/codon_start=1
/transl_table=11
/product="putative sialic acid synthase"
/protein_id="AAK91726.1"
/db_xref="GI:15193222"

gene

CDS

/translation="MKEIKIQNIISIEKAPLVVPEIGINHNSLELAKIMVDAAFSA
GAKIKHOTHIVEDEMSKAARKVIPGNAKISYIEMQKCALDYKDELAKLEYTKLGL
VYLSTPFRAGANLEDMGVSAFKIGSGECNNYPLIKHIAAFKPMIVSTGMNSTESI
KPIVKILLDNEIFEIVLHMTTNLYPTPHNLVRLNAMLKKEFSCMGLSDHTONLAC
LGAVVLGACVLRHFTDSMHRSGPDIVCSMDTNRKALKELIIQSEQAIIIRGNESKAA
KQEQVTIDFAFASVSVYSIKDKIKKEVLSDMTIWKRPGLGGISAAAEFENILGKALRDI
ENDAQLSYEDFA"
7938..9056
/gene="neuC1"
7938..9056
/gene="neuC1"
/codon_start=1
/transl_table=11
/product="putative N-acetylglucosamine-6-phosphate

gene

CDS

2-epimerase"
/protein_id="AAK91727.1"
/db_xref="GI:15193223"
/translation="MKKILFTGSRADYSKIISLMTVRVNSSEFELYIYATGMHLSKN
FGYTVKELKNGFNKFIYFINDKYQTDKALATIDGFSYANEIKLPDILVVGDR
EPLAAIVGALNNILVAHEEGEISGTDIDSLRAHISLAHILHVDEFKAKRMGLG
EDEKSFITGSPDLELNDKISLSAKRYIDINENYALLMHPVTTETTSKNOAD
NLVKALIQSNKVIYVPPNDLGFELILOSIEFEKNNPRKFLPSPSRFEVFTLLKNA
DFLGNSSCILKEALVLTAGILVGSQRNGRNLTKVNSDEILKAINTIHKQ
DLSAKLEILDSKLFEEYLOSGDFKLSLQTKVFKDIK"
9053..9718
/gene="neuA1"
9053..9718
/gene="neuA1"
/codon_start=1
/transl_table=11
/product="CMP-Neu5Ac synthetase"
/protein_id="AAK91728.1"
/db_xref="GI:15193224"
/translation="MSLAIPARGSGKIKNNLNNKPLIYYTIKAALNAKSISK
VVSDDSEILNAKQNDILKRPISLAQDDTSDKVLHALKFYKDYEDVVFLQPT
SLRTWHINEAFNLYKSNANALISVSECNDKILKAFVNCDCGLAGCNDIEYPM
RKLPTYMSNGAIYILKNEFLNPNFLOSKTKHFLMDESSLDIDCLDLAKVEQI
WKK"
9706..9900
/note="orf11; premature translational stop codon due to
variation in homopolymeric G-tract"
/codon_start=1
/transl_table=11
/product="putative acetyltransferase"
/protein_id="AAK91729.1"
/db_xref="GI:15193225"

Query Match

Best Local Similarity 2.8%; Score 82.4; DB 1; Length 11442;

Matches 213; Conservative 0; Mismatches 196; Indels 3; Gaps 1;

QY 1296 AAAGGTACCTTACATCATAGAAAGCTTTACCAATAGAGATTCACATATTCATAGAAAT 1355
DB 11399 AANTTCCTAAACTTGCATAGATTGTTTTATTTATTTTACAAAAATTTAGGAAAAATAT 11340
QY 1356 ACCTTTAGTTCTTATTTATATCCCGCTTATAACTGTGCAAAATATATATCAAGATGTGT 1415
DB 11339 GCCACAACTTCTATCATATAATCCCGCTTTTAAATCTTGCATTATTTATCAAGGGTTT 11280
QY 1416 AGATAGTCTCTTAATCAAACTGTTCGCATCTCGAGTTTGTATTTGTATACGATGTTTC 1475
DB 11279 GCAAAGTTGATAAATCAGACTTTAAAGATATTGAAATTTGATTATCATGATGATAAAG 11220
QY 1476 AACAGATAATACCTTAGAAGTGATCAATAGCTTTATGGTAATATCTAGGGTACGAT 1535
DB 11219 TAAGATAATAGTTTAAATATGTTTGTAGATTGCTTAAAGATCAAGGATAAAAT 11160
QY 1536 CAT---GTCTAAACCAATCGCGGAATAGCCTCAGCATCAAAATCAGCCGTTCTTTTTCG 1592
DB 11159 CTTTCAAAATGAGAAAAATTTAGCAGCTTTTGCAAGTAGAAATTTGGGGTTTGCATTC 11100
QY 1593 TAAGGTTATATACATCGGCGAGTTAGATTCAGATGATTCTTGAGCCTGATGAGTTGA 1652
DB 11099 TAGTCTGTGATTTATATATGTTTGTAGATAGTATGATTTTAAACACCTGATGCTTGTA 11040
QY 1653 ACTGTGTTTAAAGAAATTTTAAAGATAAAACCGCTAGCTGTGCTTTTATACC 1704
DB 11039 AATAACATTTAAAGAGATGAAAAAGGTTTGTATTGCTTTGTTTGTATGCC 10988

RESULT 12

AY044868/c

LOCUS

DEFINITION

AY044868 11455 bp DNA linear BCT 03-JAN-2002
Campylobacter jejuni heptosyltransferase I (waaC), putative
cds; lipid A biosynthesis acyltransferase (htrB), putative
two-domain glycosyltransferase, putative glycosyltransferase,
putative beta-1,4-N-acetylgalactosaminyltransferase (cgtA),

putative beta-1,3-galactosyltransferase (cgtB),
alpha-2,3-alpha-2,8-sialyltransferase (cstII), putative sialic
acid synthase (neuB), putative N-acetylglucosamine-6-phosphate
(neuCI), CMP-Neu5Ac synthetase (neuA), putative acetyltransferase,
and putative glycosyltransferase (waaV) genes, complete cds; and
heptosyltransferase II (waaF) gene, partial cds.

ACCESSION AY044868.1 GI:15430476

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

```

LKNGLYCYGFTISVCNALGTFPKYMKLYELOKPKTKCTCALIITTYNOKERLVLVD
SVKMLTFLPNEVLADGSKEDTARLEEBQKDPSPCSLKHWHQDEGFRAAKSNKAI
RVSKSEYIILIDGMILEKDFIKNHFEFAKRVILQGSRLILNEKSEELIKNDNYSI
AFSKDFKYSNFKLAKIYIKISKIDQKFTNKELIKGSKTCNNSFFKTFDKLUDGN
ENFTGWRGDESEFVARELNGKIPRLKFAIAHYIHEENSCKMLESNHQIYLDITIK
NKITSWR"
2773..3966
/note="orf"
/codon_start=1
/transl_table=11
/product="putative glycosyltransferase"
/protein_id="AAK95998.1"
/db_xref="GI:15430480"
/translation="MKKIGVPIYVNVKYLRECLDSVINOTYTNLEIMLVNDGSTDE
NSLNAKYITLKDRIILFDKNGGLSARNVGYIEPSGKYKLNKTKQHIKENSILIEP
OLDGNPNYIKYKSOAFNNEKDLNTFTYPSIDYIIFLDSNDWELNCLIECVIRM
KNVILWFDHCTYEDNIKKNKKTMEIFNFKKECIITPKREYANRALSVGRSDISFG
WNGMIDFNFKQIKLFINIINEDIHIFLAFASANKIYVLSQKLYLCRLRANSISN
HDKITRANVSEYFKDIYETEGENAKAKNYLKAASRVITALKIJEFFKDKOKSENALA
IKETFLPCYAKKALMIKKFKKDLNLKQLVLIKPFIQTKLPYDIWKFQWQIKKIFNN
KNIKN"
4015..5058
/genes="cgtA"
CDS
4015..5058
/genes="cgtA"
/codon_start=1
/transl_table=11
/product="putative
beta-1,4-N-acetylglactosaminyltransferase"
/protein_id="AAK95999.1"
/db_xref="GI:15430481"
/translation="MLFOSYFVIICLFIPRKIRHKIKTFLKNIQORDKIDSYPK
KTLAQINKYNEDLIKINAKLIGEHKGYFVNDKSKDPKSPLNPAWIRVKNEAITL
KASLESILPAQGVIGVNDCTDGSSEIILIECKQYPSFIPKYPYELIQINPKSEEN
KLYSYNVASFPKDEWLKIDVHIIIDAKKLYKSFIPRKNKYDVVSYSRVYDIHFN
DNFFLCKDNNNGILPEGDCULLINNYNLKWEVLDIRNNNNKKATQKQSFNSIHSL
QLKYKRRILFTELNNYHFFPLKHKRAQDIYKNNWISIEEFKFKYQLONLKHIEPSI
SKETLKIFLTLF"
complement(5044..5949)
/genes="cgtB"
CDS
complement(5044..5949)
/genes="cgtB"
/codon_start=1
/transl_table=11
/product="putative beta-1,3-galactosyltransferase"
/protein_id="AAK96000.1"
/db_xref="GI:15430482"
/translation="MFKSIIILPTVNVQYIARATESCINQTFKDIEIILVDDCGNDN
SINIAKESKDKRIKIIHNEKNLGLLRARYEGVKVANSPIYIMFLDPPDYLELNACEE
CIRLDEQDEVDLVFNAIVESNVISYKFPDNGSFYSKKEFVKKIIANKNLYTMWG
KLIRKLYLEAFASLRLEKDVKNMAEDVLLYPMLSQAQKIAYMNCNLYHYVPNNNS
ICNTNEVLVKNQIOLOLVNLVLRQNYILNKYCSVLYVLKYLILYIYIKYKRTKLM
VILLAKINILITLILFKYKKFLKQC"
6038..6913
/genes="cstII"
CDS
6038..6913
/genes="cstII"
/codon_start=1
/transl_table=11
/product="alpha-2,3-/alpha-2,8-sialyltransferase"
/protein_id="AAK96001.1"
/db_xref="GI:15430483"
/translation="MKKVIISGNPSLKEIDYSRLPNDFVRCNCFYEDKYYLGGK
KRAVYFNPISLFFQYVTLKHLIQNQYETELMCSNFNQAHLENQNFVKTFFDYFDDA
HLYGDFEKLQKLEAFKHEIYFNORITSGVYMCVVAIALGYKEIYLSGIDFYQNGS
SYAFDPKQKNLLKLPNKNDSHYIGHSKNTDIIKALEFLEKTYEIKVLCCLCPNSLLA
NFTELAPNLSNFIIOEKNNYTKDIIIPSEAYGKTKNINFKKIKENIYKYLKID
NLRPSDIDKHYFKG"
6914..7954
/genes="neuB1"
CDS
6914..7954
/genes="neuB1"
/codon_start=1

```

```

/transl_table=11
/product="putative sialic acid synthase"
/protein_id="AAK96002.1"
/db_xref="GI:15430484"
/translation="MKELKIQNIILISEKAPLVVPEIGINHNHGSLELAKIMVDAAFSA
GAKIILKHOITHIVEDMSKAARKVIPGNNAKISIEIYELMOKALNYKDELAKVEYTELGL
VYLSPTFSRAGANLEDMGVSAFKTSGECNNYPLIRKHAIAFKFKFVWYVSTGNNSTESI
KPTVKIILDNEIPFVLMHTNLYTPPHNLVRLNAMLLEKLFKFSFMSVGLSDHTDNLAC
LGAVALGACVLERHFTDSMHRSGPDIVCSMDTQALKELIIQSEOMAIKRGNNESKAA
KQEOVTIDFAFASVVSIIKDKKEVLSMDNIWVRPGLGGISASEFENILGKKALRDI
ENDTQLSVEDET"
7951..9069
/genes="neuCl"
CDS
7951..9069
/genes="neuCl"
/codon_start=1
/transl_table=11
/product="putative N-acetylglucosamine-6-phosphate"
/protein_id="AAK96003.1"
/db_xref="GI:15430485"
/translation="MKKILFITGTRADYSKISLMYRVONSSEFELYTPATGMHLSKN
FGYTVKLYKNGFKNIYEFINDYKYQTDKALATTIDGFSRYVNEIKPDLIYVHGDRIR
EPLAAIIVGALNNILVAHEGSEISGTIDDSLRLHAIKSLAHILHLVNDKFAKRLMQLG
EDEKSIPIIGSPDLLELNDNKISLSEAKKYVDYDINKVALLMHPVTTTETSIKNOAD
NLVKALISNKNYIYIPNDNLGFLBLLQSIEELKNRPKLEPSLRFYFTLLKNA
DFIIGNSSCILKEALYLTAGILVSGRQNGRUGNENTLKVANSDEILKAKINTIHKKO
DLFSAKLEILNSSKLFFEYLLSGEFFKLNTQKVFKDIIK"
9066..9731
/genes="neuA1"
CDS
9066..9731
/genes="neuA1"
/codon_start=1
/transl_table=11
/product="CMP-Neu5Ac synthetase"
/protein_id="AAK96004.1"
/db_xref="GI:15430486"
/translation="MSLAIIPARGGSKGIKNKLVLLNKNPLIYTIKALNTKSTSK
VYVSSDSEILNYAKSONVDILKRPISLAQDDTTSPKVLHLHAKFKYKDVVFLQPT
SPRTNIHIDEAFNLYKSNALISVSECDNKILKAFVNDYDGLAGICNDYEPFIP
ROKLPKTYMSNGAIYILKIKTEFLNPNPFLQNKTKHFLMDESSLDIDCLDLKKVEQI
WRK"
9719..9910
/note="orf11: premature translation stop due to deletion
in homopolymeric G-tract"
/codon_start=1
/transl_table=11
/product="putative acetyltransferase"
/protein_id="AAK96005.1"
/db_xref="GI:15430487"

```

```

Query Match      2.8%; Score 82.4; DB 1; Length 11455;
Best Local Similarity 51.7%; Pred. No. 0.00037;
Matches 213; Conservative 0; Mismatches 196; Indels 3; Gaps 1;

QY 1296 AAGGTACCTTACATCTATAGAAAGCTTTTACCAATAGAGATTACATATTCATAGAAAT 1355
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11409 AAATTTCTTAACCTTTCATAGAATGTGTTTATTTATTTATACAAATTTAGGAAAAATAT 11350
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1356 ACCTTTAGTTCTTATTTATATCCCGCTTATAATCGTGTCGCAAAATATATTCAGAGATGTT 1415
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11349 GCCACAACCTTCTATCATATATCCCGCTTTTAAATCTTGTGATTTTATCTTAAGAGCTTT 11290
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1416 AGATAGTGCCTTAACTCAACTGTGTGCGATCTCGAGTTTGTATTTCTTAACGATGGTTC 1475
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11289 ACAAGTTGTATAATCAGACTTTAAAGATATTGAATTTTGATCATCATGATAAAG 11230
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1476 AACAGATAATACCTTAGAAGTGATCAATAAGCTTTATGGTATAATCTTAGGGTACCGAT 1535
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11229 TAAAGATAATAGTTTAAATATGTTTGTAGATTTGCTAAAAAAGATCCAAAGATAAAAAT 11170
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1536 CAT---GTCTAAACCAATGCGGAATAGCCTCAGCATCAATGCAGCCGTTCTTTTTCG 1592
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11169 CTTTCAAAATGAAGAAAATTTAGGCACCTTTTTCAGTAAGAAATTTGGGGGTTTTTCATTC 11110
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```
/db_xref="GI:3724297"
/translation="MHKPTISSVIALTLTLLCGGSGNSGNTTTPPVKYFNVSFLDLDN
IAINTQNCQIFGYNEDOTKIVAVESAENNAQIVVHDEHGNFVVRYSNSIETNGLS
STFRQSDPQNGYVSFYVSEGRDNFVSTYSKDFIPESFIYAKNKNNSNCLSGS
GNNSPQSDIKSFHIPDENKLFPSLNFHFNOLSEPTNKLGRDVSVNSRRGRPLLL
TAYERTNQDVIQNLFSFKFRATNALSGQEIIDLQIDQRTPTWTVSVEISSAQLF
VDGQFNAPYALQMSLNTSSDKYSYSLNIDVNYLLKLGKEIKSPSPSPYKFKQHV
SGTSSITSHLSAENLDCFPNPEPTLIGCNHNNQCIKITHQTKDNTILRLIID
SELTVPSPGLPIFKOTIYSPYHNELPVVKFDSPLDLOELKEIKDFSVMKIPSPSSIT
ETFLYQDLYRTKLGSELIDFSIDNIPLLKNISAOQDQDQLLKROPYTWVLEE"
complement(10522..12714)
gene
/sequence="wbfb"
CDS
complement(10522..12714)
/sequence="wbfb"
/sequence="wbfb"
/sequence="unknown"
/codon_start=1
/transl_table=11
/protein_id="BAA33587.1"
/db_xref="GI:3724298"
/translation="MKTGHRPLNLTSLSLGLVLISSALLTPROAFADFSYPLRHQS
TDFGCTGLMOPSAARMELEGELNLAITNDDYLHYAISLQLPWLTTVRYTQVHELL
YSDQASFGSDTYTDXSIDAKLRLWQESYHLPSELSGLRDIGCTGLFDEGLVATKQV
GPDFTLFGWGTGLGNANLSDGKISSDCGSDNSYGGNGMLDILGRMTGCTALFEGG
IYQPTLAPLRLQLEYNDYRSDFPVTRGQPDAMPVTPWNFGLTYALTDWADLRVS
YQGTNLTAGLFTGNLATLKPTWLDSPATYQPVSKADLSDDEWQRLSQDVANLAG
YQVEIYQDSDSVIKGSOVKYRDREQAERAAATLLANTGINARYYRIETSDOQPLT
ETOLNSAEPKRIIDHYPNARLSDAKOHTDPAPIHGDLKAARTERWKFGFSPVLOQSF
GAEDPYLYAIGNANASYKMGDHWLQSLTYGNLIDNDYDKNYTPVPPGDTNLKRYRT
LSRQTFEELRLSHQLTYTFRWQHFTYQIAGYLETMEFAGVSGSELYLPLGKNVAWG
IDGNYVQKDRPSAFLGYEHEQYDVETGDNRYRVQGTITGTHATLYWQFQWLSLNDT
LLKVSAGRYLTEDIGFTVDFSKQFASGVVAGAFATKTDLSAEFEGSGSTKGYIISIP
LDLMTVRSERVNTSWLPLORDGGQMLSRQYGLYKMTDARSPWYERF"
complement(12711..13577)
gene
/sequence="wbfb"
CDS
complement(12711..13577)
/sequence="wbfb"
/sequence="unknown"
/codon_start=1
/transl_table=11
/protein_id="BAA33588.1"
/db_xref="GI:3724299"
/translation="MCQTDLPLRLSPFKICRSPKIKRAMSFITSRGWLTAAAIISGH
SICVAAFLATSEVRAVHNNQAIHSTRLSFSEPMRLDSAVMTQLEAGLNTQQVWPS
AGLFDLSHAFLEKRDVLLKLDQSSAPPTQOALWASLIAQLRQAEFAKRLFTSVDPD
WTRIAPOHNPRLNGSWLLTNSKSTQSVYGVAVNOPGDVIMHNLRSKDYAHAAGLID
EOISEIVLQPDGIAQKHAVYHNDQFNEVAPGAIYVPLPLKRAFFDPTVTADLNO
LVIELLRNRLPL"
complement(13577..14266)
gene
/sequence="wbfb"
CDS
complement(13577..14266)
/sequence="wbfb"
/sequence="unknown"
/codon_start=1
/transl_table=11
/protein_id="BAA33589.1"
/db_xref="GI:3724300"
/translation="MHRLYLAFCOLFLIMGCSNSSHSETMOLAIFGTPAPVISPO
KVAGIPYASLYIQONDNPLAMVLAWAEPTRNSRHRPALWVSAGQEMLYTEGGRITKT
NLMDGANLSLSDSDPLGLLHRASSTPREWHFYSWQPGYHFRYLARHSQFTLIGEQ
QKOLPNGESRTLLRIDEQITIAALNQYHNOYWLPESTGAVTASEQOOLAPGTHRYALA
VAKSYLNQGGI"
14390..14638
gene
/sequence="orf6"
CDS
14390..14638
/sequence="orf6"
/sequence="unknown"
/codon_start=1
/transl_table=11
/protein_id="BAA33590.1"
/db_xref="GI:3724301"
/translation="MATTATAAAVATAVVAVPTAAVLAVEPPALAAVSSVAVAKPL
TSSALSNTSAEFFIVISMKDIYITDVHSPFRTRTKCR"
14935..15525
```

```
/sequence="wbfb"
14935..15525
gene
/sequence="wbfb"
/sequence="unknown"
/codon_start=1
/transl_table=11
/protein_id="BAA33591.1"
/db_xref="GI:3724302"
/translation="MQOSIHASLARNVFKTSPLRFLSTPLCQCVALMTSSLLTSSLLV
PSVAAEPYAGISLAYTNAEYQPNHAASTGDSPIAVALQGGTFFSDTLGLAEARTSTV
QRSQGVNIDNLSAAFVKYKLNLPMSAQSLYGLAGYSYFSLDKQGVGVSDSGFSSGIGI
HYALDSRSIAFAIDFINHANGDEARLNTVNSLQKF"
15645..18380
gene
/sequence="wbfb"
CDS
15645..18380
/sequence="wbfb"
/sequence="probable transport of capsular precursors"
/codon_start=1
/transl_table=11
/protein_id="BAA33592.1"
/db_xref="GI:3724303"
/translation="MTQGYILRTLSSLVACLSPVAMAQTPPEQIQMFKNLPAQAQQA
LASKYGFALPSATNSASQYDNPEVMLPQNLNENITAOPLPGODALERRGLSLFSG
SPSTFAPGVDPVPSDYTVGAGDEIVQLFGKQNTQHLRYNRAGVINFTSLGVPQVA
GLKFSFVRQSLGQVRVEQMGICIRSDISLGLRMTQIFVMGDAYKPGATVTSALTLSQ
AIYSGFSQSGALNLTQLAREGLVQLDLYLLKGDNRNDVRLAMPGDVVFVIGAVD
NTVEIRGEVNRPAVYEVKSTETQDIIKMGAGFTANAYRDKTEVKRYASDGIKRALTL
NLNEGLOTKLNGDAINVLKKTQALAHAYEIVGDVHQHGFIEWKFGLRVADVQSV
DSAFNATDVNYALVVRQKNLQRDIEVQOINLANAILEPGRSDNLTQSRDRLLIFNR
FNFDYDNTDLSKDETAKTLEQAQOATQAKLEQKQOQINNTVRYNASALNPS
LPNNLSQLPSGQSLAEHPEAKIQFAGQEMTLEAEFKLQNTNRLLAPILQLLYQ
QAQGVSPQIAEVVGEVYKPYPIAAMPVSALLAAGGLTNAYTLRABLARREID
PNERYVTQISSLDLRQMLSEPSANRWTIOGMDKLVLEKPAKIQOQTVVLQGVLFPP
CTYVRQGETLAELLRAGGLTEFANPRGAIFREALRLOEQKLLNOYAADRAETAK
KTPRADNLSSVSDPAKTLAFVEASRALKRMVYVQLSRIMDEPADGPMLEDGDF
LYFTPTFTISINGEVOVPTYLLDGLNVDYDLYNKAGGVKQADADRIFFYRADGSV
18485..19492
gene
/sequence="wzz"
CDS
18485..19492
/sequence="wzz"
/sequence="probable regulation of O-antigen length"
/codon_start=1
/transl_table=11
/protein_id="BAA33593.1"
/db_xref="GI:3724304"
/translation="MQEQKUTPPNYLTPAQPTAIGDDEIDLRELFRALWKGKWTIIA
ITLVFAIGSVIFAIMQPNYKAEALIAPEAQGGGLSALASQFGGLASAGVNLGGK
GYVDKQMAIEVLKSRQTFSEFIOOHNLADINAAKKWDSADQLIYDPALYNQOHTT
WREKAPKPPPEPSMOBAYKVFSKLHVNKAKDSGMVTIAVEHLSPTVAQOQWTLIE
DINKVMKRDVAEHRSTAFLENGIALTNVADITVLYKLEEQAKTIFAEVDEYV
FKTIDPALAPEAKPKRALICVGI.LGMLGVAVILVRFAPFRKAGNHEHADVRET
A"
19650..20777
gene
/sequence="IS1358"
CDS
19650..20777
/sequence="IS1358"
/sequence="probable insertion sequence like element"
/codon_start=1
/transl_table=11
/protein_id="BAA33594.1"
/db_xref="GI:3724305"
/translation="MSELINPEMHFOIKDYROBSKVEHKLSDIILLTCGVLSCHDG
WDGIIDFGNARLDFLKRYGHFEAGIPSDATLSRVGMGINPVALORSFIAMKNCHTILT
DGEVIAIDGKTRLSYDRSKGKTIHMVNAFATANGSIGOLKYSNEITAIPLKLL
DLIDVKGCLITIDAMGCKKIAQRIKADYLLAVKNGQMLQFQAFDDYFRMDMLQD
FDGSSYSTQKSHGRIETRYALVNRDLSVLGDIIEHWPGLKSMGIVASIROESAVATE
QDVSIRYICSKLEAQTLLLEATRSHMGVEVHWSLDTAFCDNSRTRADDAEAFAR
IROICLNLLKSETTFKGIKRRKNCAMDEKYLKSVLESLT"
20961..22079
gene
/sequence="gmd"
CDS
20961..22079
/sequence="gmd"
```



```
LLSLAKTIPTIGYAIWGTIGAGVSVTLGMIVFKERKSVGKLLFTITMIAGVVGKLK
TSGV"
/Join(2650..3774,3788..3822)
/gene="ddla"
RBS
2650..2655
/gene="ddla"
CDS
2662..3774
/gene="ddla"
/codon_start=1
/transl_table=11
/product="D-alanine-D-alanine ligase"
/protein_id="CAC98933.1"
/db_xref="GI:16410258"
/db_xref="SPTREMBL:Q8Y8P1"
/translation="MKTKLILLYGKSAEHSVLSQAFSVINALDLKREAPAIYITN
EGEMTQGLPSKLDPEVOLRESATDTIKLATSEKSEGEALSPAVLEADGOETVFE
RLIHGPNGEDGVQGLFEVLNIPDPVGVNGVLASSAAMDKIVMKKIFADGIPQPAVAV
RLIDMKNIQEEWVAEMEEVLTYPVFVKPANLGSSVIGISKATNKKKELVADWAEFLYDR
RVVVGQVVAIEHTEHGVLDNDTPVCSVGEILPEGAVALFYDYKAKYQDNNTALIIPT
EYDPELEQMKRYAIOAFLGLDASGLVRAFDFTEDNQLFLNEVNTMPGFTPYSMYPL
LWQETGLPYGALIERLVDLAKERHAAKNALKYKLED"
/Join(3822..3828,3838..5211)
/gene="ddla"
/Join(3822..3828,3838..5211)
/gene="murF"
RBS
3822..3828
/gene="murF"
CDS
3838..5211
/gene="murF"
/codon_start=1
/transl_table=11
/product="UDP-N-acetylmutramoylalanyl-D-glutamyl-2,
6-diamino pimelate-D-alanyl-D-alanyl ligase"
/protein_id="CAC98934.1"
/db_xref="GI:16410259"
/db_xref="SPTREMBL:Q8Y8P0"
/translation="MKTKVGEVVMALDSSIHVDAFRDVVITGVCFDTRQIKSGDLFVP
FYGNVRDHFEGVQAREMGAVATFWQKNVPNPPTDFPVLIVEDTLALQELAAKYIQQ
VKPKVIAITGSKGKTTKDIMAIAIVETTVKHYTGNGFNHNGHGLPTILTMPEDTEVA
VLEMGMRHEIEVLISKIAKPDIAITNTGEAHEIYVLSREIEIAKAKLETACGLNPSG
ILYIHEETVLLGNTNGDFRQLTFCKSEAAEYPLAIRAEAGTSFTWNWEPELEIFV
PIIHEHNTWAAMLAAREIGIEGKESQALKNMERSKSRLEWITKSGARILNDAY
NYSPTALKTVLFTFMHMDSSGPKYLVLDMLDELGLDLTKLHQSVAEVLNDTIQKVF
LXGDAMKAFGEVAESKIGQGKVHHPDFTKESLETALLSEIKGNENWILVKGSGMGLKDV
VENLIK"
/Join(5263..5268,5278..5991)
/gene="lmo0857"
RBS
5263..5268
/gene="lmo0857"
CDS
5278..5991
/gene="lmo0857"
/note="similar to carboxylesterase"
/codon_start=1
/transl_table=11
/protein_id="CAC98935.1"
/db_xref="GI:16410260"
/db_xref="SPTREMBL:Q8Y8N9"
/translation="WVACLCIHGFTGSPSEVKPLADYLRHETDWDVLAPTLPCHDHUR
HLKNVYKDWIVFVDSILSOMLKEDDEVYIIGFSGMGLLAGWLARHYPEVKKLVLLST
AVNMEPVLVNSQKVLTEAKEVTLKNSPMFKRYQKVTETPTWTLQFQKRWALAK
PVFEHIEIPTFAQGSADQVVPAAKSVNFMESIPGPKELFILEGSKHVICQDEQADK
LFAAVLTLQKDVAVLNTK"
complement(Join(5990..7028,7035..7040))
/gene="lmo0858"
terminator
complement(5990..6032)
/gene="lmo0858"
RBS
5998..6036
/gene="lmo0857"
terminator
5998..6036
/gene="lmo0857"
complement(6003..7028)
/gene="lmo0858"
/note="similar to transcription regulator lacI family"
```

```
/codon_start=1
/transl_table=11
/protein_id="CAC98936.1"
/db_xref="GI:16410261"
/db_xref="SPTREMBL:Q8Y8N8"
/translation="MKTKMODIADRLNITKNVSQALGNKNKNGVSEOTKLAFFKMADE
LGKYYKREIAREVVKERFALLATSFALSQRSFFGEIIDNMKQSIHAHQAEIIFPVTD
EEAATNOLPQLTSENWTFILSHINTDYSKVIDILMPVVLIDHDPHLKADAVIS
ONKDGAFMAVEFLIONKHOKIGPLGDVTFSPYERLEGYKKALQYXHIPNKKYAIT
RIKEEOTLYRTLDELDELPTAFVNSGLGFIINTVYLOSQGYNIPKDLSLIICFDNTE
FTVLSNQLFTTCTNLNPFMGEKAVELMYNRIRKPDGDFVHLALATNLRIARSVGENKK
Query Match 2.8% Score 82.4; DB 1; Length 270050;
Best Local Similarity 55.0%; Pred. No. 0.00017;
Matches 186; Conservative 0; Mismatches 146; Indels 6; Gaps 1;
Qy 1345 ATTCATAGAATACCTTTAGTTTCTTATTTATATCCCGCTTATAACTGTGCAAAATATATT 1404
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 234478 ATTCTTTTGGAAACCTTTAGTAGTGTAAATTTATCCAGTTTATAATGTAGAAAGATGTT 234537
Qy 1405 CAAAGATGCTGTAGTAGTGCTCTTAATCAAACTGTTGTCGATCTCGAGTTTGTATTCT 1464
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 234538 AACGATGCTTGGATTCAGTTCTTTGAGCAACATATACCTCGAAGTTATCGTAGTT 234597
Qy 1465 AACGATGCTTCAACAGATAATACCTTTAGAAAGTGAATCAAGCTTTATGGTAATAATCCT 1524
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 234598 AATGATGAGCGACTGATAATTACGTAAGTCATAAAAAGTATCTCAGATAA-----T 234651
Qy 1525 AGGTACGCATCATGCTCTAAACCAAAATGCGGAATAGCCTCAGCATCAAAATCGAGCCGTT 1584
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 234652 CGAATTCGTTATTTTGAAGAAAGAAAATGCGGACAAGTACTGCCCGAAATTTTGGATTA 234711
Qy 1585 TCCTTTGCTAAAGGTTATTACATTTGGCGAGTTTAGATTTCAGATGATTATCTTCAGCCTGAT 1644
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 234712 GATGTGGCTACGGGAGATTACATCGTGATGGTAGATAGTGATGATTACATTAGCAAAAAC 234771
Qy 1645 GCAGTTGAACGTGTGTTTAAAAAGAAATTTTAAAAAGATAA 1682
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 234772 CTAGTAGAAACTTGTGTTTGCATACTCTACAAAAAACGAA 234809
```

Search completed: January 3, 2003, 23:56:07
Job time : 8776 secs

THIS PAGE BLANK (USPTO)

BEST AVAILABLE COPY

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2003, 20:10:07 ; Search time 574 Seconds
(without alignments)

11687.636 Million cell updates/sec

Title: US-09-842-484A-1

Perfect score: 2979

Sequence: 1 ttataaacgattaaagaag.....acattgcattttattaaaa 2979

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : -

N_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2979	100.0	2979	21	AAA27449
2	2979	100.0	2979	24	ABA05097
3	2963	99.5	2979	24	ABA05098
4	2283.4	76.6	2937	20	AA558857
5	2283.4	76.6	2937	21	AA235589
6	1503.2	50.5	2112	21	AAA27448
7	86	2.9	5550	20	AA13074
8	81	2.7	984	24	ABQ68041
9	81	2.7	984	24	ABQ69955

10	80	2.7	19966	21	AA245258
11	80	2.7	19966	21	AA230355
12	79.2	2.7	11474	21	AA53720
13	77.2	2.6	12237	24	ABL34358
14	77.2	2.6	17527	24	ABL33432
15	77.2	2.6	17527	24	AA563332
16	76	2.6	7458	21	AAA70106
17	75.6	2.5	963	24	ABN67436
18	74.8	2.5	15548	24	ABL34155
19	74.6	2.5	13807	18	AAT73236
20	74.6	2.5	14602	18	AAT73236
21	74	2.5	40862	24	ABL34072
22	73.2	2.5	6992	21	AA260930
23	73	2.5	1082	24	ABQ69025
24	73	2.5	1162	24	ABQ70597
25	71.8	2.4	9106	24	ABL54369
26	71.8	2.4	17280	22	AAS46772
27	70.8	2.4	981	24	ABN67435
28	70.6	2.4	5748	24	ABL33142
29	70	2.3	14924	24	ABL54321
30	70	2.3	14924	24	ABL32224
31	69.8	2.3	6048	24	ABQ67002
32	69.8	2.3	16258	24	ABL70376
33	69.8	2.3	16258	24	ABK40038
34	69.4	2.3	7442	22	AAS46686
35	68.8	2.3	6109	24	ABL32326
36	68.8	2.3	6109	24	AA561077
37	68.8	2.3	9095	24	ABQ67061
38	68.8	2.3	12507	24	ABL32299
39	68	2.3	5181	24	ABL70443
40	68	2.3	5493	24	AA561386
41	68	2.3	6244	24	ABL32484
42	68	2.3	15950	21	AA245260
43	68	2.3	15950	21	AA230357
44	67.8	2.3	2325	24	ABQ67858
45	67.8	2.3	2325	24	ABQ69470

ALIGNMENTS

RESULT 1
AAA27449
ID AAA27449 standard; cDNA; 2979 BP.
XX
AC AAA27449;
XX
DT 17-AUG-2000 (first entry)
XX
DE P. multocida chondroitin synthase coding sequence.
XX
KW Chondroitin synthase; CS; enzyme; hyaluronic acid; ulcer;
KW tissue abrasion; viscoelastic replacement; bioadhesive;
KW ss.
XX
OS Pasteurella multocida.
XX
FH Key Location/Qualifiers
FT CDS 61..2958
FT FT /*tag= a
FT FT /product= "PmCS"
XX
PN WO200027437-A2.
XX
PD 18-MAY-2000.
XX
PF 10-NOV-1999; 99WO-US26501.
XX
PR 11-NOV-1998; 98US-0107929.
PR 01-APR-1999; 99US-0283402.
XX
PA (OKLA) UNIV OKLAHOMA STATE.
XX


```
Db 1741 GCTAATGGTTACAAATGGCCAGAAATTTTACAGAGAAAAAATCTACAACGGCTATGATGCT 1800
QY 1801 CACATTTTGAATGTTTACGATTTAGAGCTTGGCATTTAAACGGATGGATTTAAACGAAAT 1850
Db 1801 CACATTTTGAATGTTTACGATTTAGAGCTTGGCATTTAAACGGATGGATTTAAACGAAAT 1860
QY 1861 ATTGAAACCCGCGGATTTATGACATGTTCTTTAACTCAGTGAAGTTGGAAAAATTTAAA 1920
Db 1861 ATTGAAACCCGCGGATTTATGACATGTTCTTTAACTCAGTGAAGTTGGAAAAATTTAAA 1920
QY 1921 CATCTTAATAAATCTGCTATACCCGGGTATTACATGGTGATACACATCCATTAAGAAA 1980
Db 1921 CATCTTAATAAATCTGCTATACCCGGGTATTACATGGTGATACACATCCATTAAGAAA 1980
QY 1981 CTGGCATTCAAAAGAAAAACCATTTTGTGTAGTCAATCAGTCATTAATAATAGACAAGGC 2040
Db 1981 CTGGCATTCAAAAGAAAAACCATTTTGTGTAGTCAATCAGTCATTAATAATAGACAAGGC 2040
QY 2041 ATCAATTTATTATTAATATGACAAATTTGATGATTTAGATGAAGTGAAGATATATCTTC 2100
Db 2041 ATCAATTTATTATTAATATGACAAATTTGATGATTTAGATGAAGTGAAGATATATCTTC 2100
QY 2101 AATAAAACCCGCGATATCAAGAAGAAATGGATATGTTAAAGATCTTAAACTCATTCAA 2160
Db 2101 AATAAAACCCGCGATATCAAGAAGAAATGGATATGTTAAAGATCTTAAACTCATTCAA 2160
QY 2161 AATAAGATGCAAAATCGCAGTCAGTATTTCTATCCCAATACATTAACGGCTTAGTG 2220
Db 2161 AATAAGATGCAAAATCGCAGTCAGTATTTCTATCCCAATACATTAACGGCTTAGTG 2220
QY 2221 AAAAACTAACCAATATTATTGAATATAATAAAAAATATTTCTGTTATTTCTACATGTT 2280
Db 2221 AAAAACTAACCAATATTATTGAATATAATAAAAAATATTTCTGTTATTTCTACATGTT 2280
QY 2281 GATGAATCATCTTACACAGACATCAAAAAGAAATATTTGGCTTTCTATCATAAGCAC 2340
Db 2281 GATGAATCATCTTACACAGACATCAAAAAGAAATATTTGGCTTTCTATCATAAGCAC 2340
QY 2341 CAAGTGAATATTTTACTAAATATGACATCTCATATTACAGAGTAAATAGACTAATAAAA 2400
Db 2341 CAAGTGAATATTTTACTAAATATGACATCTCATATTACAGAGTAAATAGACTAATAAAA 2400
QY 2401 ACTGAGCACATTTAAGTAAATATAATAAATAAGTCAGTTAAATCTAAATTTGTGAATAC 2460
Db 2401 ACTGAGCACATTTAAGTAAATATAATAAATAAGTCAGTTAAATCTAAATTTGTGAATAC 2460
QY 2461 ATCAATTTTGAATATGACAGCCTATTCGTTAAAAATGACAGCTATGCTTATATGAAA 2520
Db 2461 ATCAATTTTGAATATGACAGCCTATTCGTTAAAAATGACAGCTATGCTTATATGAAA 2520
QY 2521 AAATATGATGTCGGCATGAATTTCTCAGCATTAACACATGATTCGATCGAGAAAATCAAT 2580
Db 2521 AAATATGATGTCGGCATGAATTTCTCAGCATTAACACATGATTCGATCGAGAAAATCAAT 2580
QY 2581 GCGCATCCACCATTTAAAAAGCTGATTTAAACCTTATTTAATGACAATGACTTAAAGAGT 2640
Db 2581 GCGCATCCACCATTTAAAAAGCTGATTTAAACCTTATTTAATGACAATGACTTAAAGAGT 2640
QY 2641 ATGAATGTGAAGGGGATCACAAGGTATGTTTATGAAGTATGGCTACCGCATGAGCTT 2700
Db 2641 ATGAATGTGAAGGGGATCACAAGGTATGTTTATGAAGTATGGCTACCGCATGAGCTT 2700
QY 2701 CTGACGATTTAAAGAAGTCATCACATCTCGCCCAATCAATTTGATGTGCCAGATAT 2760
Db 2701 CTGACGATTTAAAGAAGTCATCACATCTCGCCCAATCAATTTGATGTGCCAGATAT 2760
QY 2761 AACACTGAGGATATTTGGTTCCAAATTTGCACTTTTATCTTTAGAAAAAGAAACCGGCAT 2820
Db 2761 AACACTGAGGATATTTGGTTCCAAATTTGCACTTTTATCTTTAGAAAAAGAAACCGGCAT 2820
QY 2821 GTATTTTAAACATCGACCCCTGACCTTATATGCTTTGGGACGAAATTTACAATGGACA 2880
Db 2821 GTATTTTAAACATCGACCCCTGACCTTATATGCTTTGGGACGAAATTTACAATGGACA 2880
```

```
QY 2881 AATGAACAAATTTCAAAGTGCAGAAAAAGCGAAAAATATCCCGTTAAACAAGTTTCATTATT 2940
Db 2881 AATGAACAAATTTCAAAGTGCAGAAAAAGCGAAAAATATATCCCGTTAAACAAGTTTCATTATT 2940
QY 2941 AATAGTATAACGCTATATAAAACATTTGCAATTTTATTAAAA 2979
Db 2941 AATAGTATAACGCTATATAAAACATTTGCAATTTTATTAAAA 2979
```

RESULT 2

ABA05097

ID ABA05097 standard; DNA; 2979 BP.

XX ABA05097;

XX 22-FEB-2002 (first entry)

XX Pasteurella multocida chondroitin synthase gene #1.

XX Chondroitin synthase; dermatan sulphate; chondroitin sulphate polymer;
eye application; joint application; moisturiser; drug delivery;
wound dressing; biocompatible film; ds.

XX Pasteurella multocida.

XX Key Location/Qualifiers

XX CDS 61..2945

XX FT /*tag= a

XX FT /product= "chondroitin synthase"

XX WO200180810-A2.

XX 01-NOV-2001.

XX 25-APR-2001; 2001WO-US13395.

XX 25-APR-2000; 2000US-199538P.

XX (DANG/) DE ANGELIS P L.

XX De Angelis PL;

XX WPI: 2002-049237/06.

XX P-PSDB: RAM47335.

XX New chondroitin synthase gene obtained from Pasteurella multocida,
useful as hyaluronan polysaccharide substitute in medical or cosmetic
applications, e.g. for eye or joint applications, for moisturizer or
wound dressings

XX Claim 4; Page 117-118; 125pp; English.

XX The present invention relates to the coding sequence of the Pasteurella
multocida chondroitin synthase. A chondroitin polysaccharide may be used
as a hyaluronan polysaccharide substitute in medical or cosmetic
applications, for example in eye or joint applications, for moisturiser
or wound dressings. The enzyme may be used in covalently coupling
specific drugs, proteins or toxins to the structurally modified
chondroitin for general or targeted drug delivery or radiological
procedures, covalently cross linking the hyaluronic acid itself or to
other supports to achieve a gel or other three dimensional biomaterial
with stronger physical properties, and covalently linking hyaluronic acid
to a surface to create a biocompatible film or monolayer. The present
sequence is one version of the coding sequence of the invention.

XX SQ Sequence 2979 BP; 1129 A; 466 C; 497 G; 887 T; 0 other;

Query Match

Best Local Similarity 100.0%; Score 2979; DB 24; Length 2979;

Matches 2979; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTATAAAGTGAATTAAGAAGGTAAACGATTCAAGCAAGGTTAATTTTAAAGGAAGAAA 60

|||||
Db 1 TTAATAACTGATTAAGAGGTAAACGATTCAAGCAAGTTAAATTTTAAAGGAAGAA 60
QY 61 ATGAATACATATTACAAAGCAATAAAAGCATATAACAGCAATGACTATGAATTAGCACTC 120
|||||
Db 61 ATGAATACATATTACAAAGCAATAAAAGCATATAACAGCAATGACTATGAATTAGCACTC 120
QY 121 AAATTTATTGAGAAGCTGCTGAACCTTACGGCGGAAATAATCGTTGAATTCCAAATATTC 180
|||||
Db 121 AAATTTATTGAGAAGCTGCTGAACCTTACGGCGGAAATAATCGTTGAATTCCAAATATTC 180
QY 181 AAATGTAAGAAAAAATCTCGACCAATTTCTTATGTAAGTGAAGATAAAAAACAGTGT 240
|||||
Db 181 AAATGTAAGAAAAAATCTCGACCAATTTCTTATGTAAGTGAAGATAAAAAACAGTGT 240
QY 241 TCGGATAGCTCAATAGATATCCACACAGCTCTTACTTCCAAACGTAAAAAATAACT 300
|||||
Db 241 TCGGATAGCTCAATAGATATCCACACAGCTCTTACTTCCAAACGTAAAAAATAACT 300
QY 301 CTATCGGAATCAGAAAAACAGTTTAAAAATAAATGGAATCTATCACTGGGAAAAAA 360
|||||
Db 301 CTATCGGAATCAGAAAAACAGTTTAAAAATAAATGGAATCTATCACTGGGAAAAAA 360
QY 361 TCGGAGAACCGAGAAATCAGAAAGGTGGAACCTAGTACCCAAAGATTTTCTCTAAAGATCTT 420
|||||
Db 361 TCGGAGAACCGAGAAATCAGAAAGGTGGAACCTAGTACCCAAAGATTTTCTCTAAAGATCTT 420
QY 421 GTTCTGTGCTTCCATTGCCAGATCATGTTAATGATTTTACATGGTACAAAAATCGAAAAAA 480
|||||
Db 421 GTTCTGTGCTTCCATTGCCAGATCATGTTAATGATTTTACATGGTACAAAAATCGAAAAAA 480
QY 481 AGCTTAGGTATAAGGCTGTAATAGAAATATCGGTCTTTCTTATTTATTTCTCTACATTT 540
|||||
Db 481 AGCTTAGGTATAAGGCTGTAATAGAAATATCGGTCTTTCTTATTTATTTCTCTACATTT 540
QY 541 AATCGTAGCGGTATTTAGATATAAGCTTAGCTGAGCTGTTGGTCAATCAGAAAAACAACCTAC 600
|||||
Db 541 AATCGTAGCGGTATTTAGATATAAGCTTAGCTGAGCTGTTGGTCAATCAGAAAAACAACCTAC 600
QY 601 CCATTTGAAGTCTGCTTGCAGATGATGTTAGTAAAGAAAAAATCTTACTTCCATTGTGCAA 660
|||||
Db 601 CCATTTGAAGTCTGCTTGCAGATGATGTTAGTAAAGAAAAAATCTTACTTACCATTGTGCAA 660
QY 661 AAATAGCAACAAAACTTGACATAAAGTATGTAAGACAAAAAGATTATGGATATCAATTG 720
|||||
Db 661 AAATAGCAACAAAACTTGACATAAAGTATGTAAGACAAAAAGATTATGGATATCAATTG 720
QY 721 TGTGCACTCAGAACTTAGTGTACGACGAAAGTATGATTTGTCTCGATTCTAGAC 780
|||||
Db 721 TGTGCACTCAGAACTTAGTGTACGACGAAAGTATGATTTGTCTCGATTCTAGAC 780
QY 781 TCGGATATGSCACCAACAATAATATGGTTCTATTCTTACAGAACTATTAGAAAGAC 840
|||||
Db 781 TCGGATATGSCACCAACAATAATATGGTTCTATTCTTACAGAACTATTAGAAAGAC 840
QY 841 AATGATATTGTTTAAATTTGACCTAGAAAAATATGTGGATACTCATATATATACCGCAGAA 900
|||||
Db 841 AATGATATTGTTTAAATTTGACCTAGAAAAATATGTGGATACTCATATATATACCGCAGAA 900
QY 901 CAATTCCTTAACGATCCATATTTAATAGAAATCACTACCTGAAACCGCTACAAATAACAAT 960
|||||
Db 901 CAATTCCTTAACGATCCATATTTAATAGAAATCACTACCTGAAACCGCTACAAATAACAAT 960
QY 961 CCTTCGATTACATCAAAAGAAATATATCGTTGGATTGGAGATTAGAAACATTTCAAAAAA 1020
|||||
Db 961 CCTTCGATTACATCAAAAGAAATATATCGTTGGATTGGAGATTAGAAACATTTCAAAAAA 1020
QY 1021 ACCGATAATCTAGCTCTATGATTTCCGTTTCGTTTATTTGTTGCGGGAATGTTGCA 1080
|||||
Db 1021 ACCGATAATCTAGCTCTATGATTTCCGTTTCGTTTATTTGTTGCGGGAATGTTGCA 1080
QY 1081 TTTTCTAAAGAAATGGCTTAAATAAGTAGTGTGTTTCCATGAGAAATTTAATCATTTGGGG 1140
|||||

Db 1081 TTTTCTAAAGAAATGGCTTAAATAAGTAGTGTGTTTCGATGAAGAAATTTAATCATTTGGGG 1140
QY 1141 GCGGAAGATGTAGAATTTGGTTACAGATTATTTGCAAGAGCTGTTTTTTCAGAGTAAT 1200
|||||
Db 1141 GCGGAAGATGTAGAATTTGGTTACAGATTATTTGCAAGAGCTGTTTTTTCAGAGTAAT 1200
QY 1201 GACGCGGAATGGCCATCCATCAAGAACCCCTGGTTAAAGAAAAATGAAGAACGCGAA 1260
|||||
Db 1201 GACGCGGAATGGCCATCCATCAAGAACCCCTGGTTAAAGAAAAATGAAGAACGCGAA 1260
QY 1261 GCTGTTAAAGTATTACGCTTAAATTTGTAAGAAAAAGTACCTTACATCTATAGAAAG 1320
|||||
Db 1261 GCTGTTAAAGTATTACGCTTAAATTTGTAAGAAAAAGTACCTTACATCTATAGAAAG 1320
QY 1321 CTTTACCATAAGAGATTACATATTTATAGAAATACCTTTAGTTTCTATTTATATATCC 1380
|||||
Db 1321 CTTTACCATAAGAGATTACATATTTATAGAAATACCTTTAGTTTCTATTTATATATCC 1380
QY 1381 GCTTATAACTGTGCAAAATTTATTTCAAAGATGTAGATAGTCTCTTAAATCAAACTGTT 1440
|||||
Db 1381 GCTTATAACTGTGCAAAATTTATTTCAAAGATGTAGATAGTCTCTTAAATCAAACTGTT 1440
QY 1441 GTGATCTCAGGTTTGTATTTGTAACGATGGTTCAACAGATAATACCTTAGAAGTGATC 1500
|||||
Db 1441 GTGATCTCAGGTTTGTATTTGTAACGATGGTTCAACAGATAATACCTTAGAAGTGATC 1500
QY 1501 AATAAGCTTTATGGTAAATAATCCTAGGGTACGATCATGTCTAAACCAATGCGGAATA 1560
|||||
Db 1501 AATAAGCTTTATGGTAAATAATCCTAGGGTACGATCATGTCTAAACCAATGCGGAATA 1560
QY 1561 GCTCAGCATCAATCGACGCCGTTTCTTTGCTTAAAGGTTATTACATTTGGCGAGTTAGAT 1620
|||||
Db 1561 GCTCAGCATCAATCGACGCCGTTTCTTTGCTTAAAGGTTATTACATTTGGCGAGTTAGAT 1620
QY 1621 TCAGATGATTATCTTGAGCCTGATGCACTTGAACCTGTGTTTAAAGAAATTTTAAAGAT 1680
|||||
Db 1621 TCAGATGATTATCTTGAGCCTGATGCACTTGAACCTGTGTTTAAAGAAATTTTAAAGAT 1680
QY 1681 AAAACGCTAGCTTGTGTTTATACACTAATAGAAACGTCATCCGGATGCTAGCTTAATC 1740
|||||
Db 1681 AAAACGCTAGCTTGTGTTTATACACTAATAGAAACGTCATCCGGATGCTAGCTTAATC 1740
QY 1741 GCTAATGGTTTACAAATTTGCCAGAAATTTTACAGAGAAAACTCACAACGGCTATGATTGCT 1800
|||||
Db 1741 GCTAATGGTTTACAAATTTGCCAGAAATTTTACAGAGAAAACTCACAACGGCTATGATTGCT 1800
QY 1801 CACCATTTTAAATTTTACGATTAGAGCTTGGCATTTAAACGGATGATTAAACGAAAT 1860
|||||
Db 1801 CACCATTTTAAATTTTACGATTAGAGCTTGGCATTTAAACGGATGATTAAACGAAAT 1860
QY 1861 ATTGAAACCGCTGGATTATGACATGTTCCTTAAACTCAGTGAAGTTGGAATTTTAA 1920
|||||
Db 1861 ATTGAAACCGCTGGATTATGACATGTTCCTTAAACTCAGTGAAGTTGGAATTTTAA 1920
QY 1921 CATCTTAATAAATCTGCTATAACCGGTTTACATGGTGATAACACATCCATTTAAAGAA 1980
|||||
Db 1921 CATCTTAATAAATCTGCTATAACCGGTTTACATGGTGATAACACATCCATTTAAAGAA 1980
QY 1981 CTGGCATTTCAAAAGAAAAACCATTTTGTGTAGTCAATCAGTCAATTAATAGCAAGGC 2040
|||||
Db 1981 CTGGCATTTCAAAAGAAAAACCATTTTGTGTAGTCAATCAGTCAATTAATAGCAAGGC 2040
QY 2041 ATCAATTTATTAATTTATGACAAATTTTAGATGAAAGTAGAAGTATATCTTC 2100
|||||
Db 2041 ATCAATTTATTAATTTATGACAAATTTTAGATGAAAGTAGAAGTATATCTTC 2100
QY 2101 AATAAAACCGCTGAATATCAAGAAAGAAATGGATATGTTTAAAGATCTTAAACTCATTC 2160
|||||
Db 2101 AATAAAACCGCTGAATATCAAGAAAGAAATGGATATGTTTAAAGATCTTAAACTCATTC 2160
QY 2161 AATAAAGATGCAAAATCGCAGTCAGTATTTTCTATCCCAATACATTTAAACGGCTTAGTG 2220
|||||
Db 2161 AATAAAGATGCAAAATCGCAGTCAGTATTTTCTATCCCAATACATTTAAACGGCTTAGTG 2220

```
Qy 2221 AAAAACTAAACAATATTATTGAATATAATAAAAAATATATTCGTTATTATTCTACATGTT 2280
Dy 2221 AAAAACTAAACAATATTATTGAATATAATAAAAAATATATTCGTTATTATTCTACATGTT 2280
Qy 2281 GATAAGATCATCTTTACACAGACATCAAAAAAGAAATATTGGCTTTCTATCATAAGCAC 2340
Dy 2281 GATAAGATCATCTTTACACAGACATCAAAAAAGAAATATTGGCTTTCTATCATAAGCAC 2340
Qy 2341 CAAGTGAATATTTACTATAAATTAAGTCAATCTCATATTTACAGAGTAATAGACTAATAAAA 2400
Dy 2341 CAAGTGAATATTTACTATAAATTAAGTCAATCTCATATTTACAGAGTAATAGACTAATAAAA 2400
Qy 2401 ACTGAGGCACATTTAAAGTAATTAATAAATTAAGTCAAGTTAAATCTAAATTTGTGAATAC 2460
Dy 2401 ACTGAGGCACATTTAAAGTAATTAATAAATTAAGTCAAGTTAAATCTAAATTTGTGAATAC 2460
Qy 2461 ATCAATTTTGTATATCATGACAGCCTATTCGTTAAAAATGACAGCTATGCTTTATATGAAA 2520
Dy 2461 ATCAATTTTGTATATCATGACAGCCTATTCGTTAAAAATGACAGCTATGCTTTATATGAAA 2520
Qy 2521 AAATATGATGTCGGCATCAATTTCTCAGCATTAACACATGATTGGATCGAGAAAATCAAT 2580
Dy 2521 AAATATGATGTCGGCATCAATTTCTCAGCATTAACACATGATTGGATCGAGAAAATCAAT 2580
Qy 2581 GCGCATCCACCATTTAAAAAGCTGATTAACACCTATTTTAATGACAATGACTTAAAGAACT 2640
Dy 2581 GCGCATCCACCATTTAAAAAGCTGATTAACACCTATTTTAATGACAATGACTTAAAGAACT 2640
Qy 2641 ATGAATGTGAAGGGGCATCAAGGTATGTTTATGAAGTATGCGCTACCGCATGAGCTT 2700
Dy 2641 ATGAATGTGAAGGGGCATCAAGGTATGTTTATGAAGTATGCGCTACCGCATGAGCTT 2700
Qy 2701 CTGAGGATTAATAAGAGTATCATACATCTCGCCAAATCAATGATAGTGCCAGAAATAT 2760
Dy 2701 CTGAGGATTAATAAGAGTATCATACATCTCGCCAAATCAATGATAGTGCCAGAAATAT 2760
Qy 2761 AACACTGAGGATATTTGGTTCCAAATTTGCACATTTTAACTTTAGAAAAGAAAACCGGCAT 2820
Dy 2761 AACACTGAGGATATTTGGTTCCAAATTTGCACATTTTAACTTTAGAAAAGAAAACCGGCAT 2820
Qy 2821 GTATTTAATAAACATCGACCTGACTATATATGCTTGGGAACGAAATATCAATGGACA 2880
Dy 2821 GTATTTAATAAACATCGACCTGACTATATATGCTTGGGAACGAAATATCAATGGACA 2880
Qy 2881 AATGAACAAATTCAAAGTGCRAAAAAGCGGAATATATCCCGTTTAAACAGTTTCATTAT 2940
Dy 2881 AATGAACAAATTCAAAGTGCRAAAAAGCGGAATATATCCCGTTTAAACAGTTTCATTAT 2940
Qy 2941 AATAGTATACGCTATAAAACATTTTGCATTTTATTAAAAA 2979
Dy 2941 AATAGTATACGCTATAAAACATTTTGCATTTTATTAAAAA 2979
```

RESULT 3

ABA05098

ID ABA05098 standard; DNA; 2979 Bp.

XX XX

AC ABA05098;

DT 22-FEB-2002 (first entry)

DE Pasteurella multocida chondroitin synthase gene #2.

KW Chondroitin synthase; dermatan sulphate; chondroitin sulphate polymer;
eye application; joint application; moisturiser; drug delivery;
wound dressing; biocompatible film; ds.

OS Pasteurella multocida.

PH key Location/Qualifiers
CDS 61..2958
/*tag= a

/product= "chondroitin synthase"

WO200180810-A2.

01-NOV-2001:

25-APR-2001; 2001WO-US13395.

25-APR-2000; 2000US-199538P.

(DANG/) DE ANGELIS P L.

De Angelis PL;

WPI; 2002-049237/06.

P-PSDB; AAM47336.

New chondroitin synthase gene obtained from Pasteurella multocida, useful as hyaluronan polysaccharide substitute in medical or cosmetic applications, e.g. for eye or joint applications, for moisturizer or wound dressings

Claim 4; Page 120-121; 125pp; English.

The present invention relates to the coding sequence of the Pasteurella multocida chondroitin synthase. A chondroitin polysaccharide may be used as a hyaluronan polysaccharide substitute in medical or cosmetic applications, for example in eye or joint applications, for moisturiser or wound dressings. The enzyme may be used in covalently coupling specific drugs, proteins or toxins to the structurally modified chondroitin for general or targeted drug delivery or radiological procedures, covalently cross linking the hyaluronic acid itself or to other supports to achieve a gel or other three dimensional biomaterial with stronger physical properties, and covalently linking hyaluronic acid to a surface to create a biocompatible film or monolayer. The present sequence is one version of the coding sequence of the invention.

SQ Sequence 2979 Bp; 1130 A; 466 C; 495 G; 888 T; 0 other;

Query Match 99.5%; Score 2963; DB 24; Length 2979;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2969; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 TTATAAATCGATTAAAGAGGTAACGATTCAAGCAAGGTTAATTTTAAAGGAAGAAA 60

Dy 1 TTATAAATCGATTAAAGAGGTAACGATTCAAGCAAGGTTAATTTTAAAGGAAGAAA 60

Qy 61 ATGAATACATTTATCACAAGCAATAAAGCATATATAACAGCAATGACTATGAATTTAGCACTC 120

Dy 61 ATGAATACATTTATCACAAGCAATAAAGCATATATAACAGCAATGACTATGAATTTAGCACTC 120

Qy 121 AAATTTATTTGAGAGTCTGCTGAAACCTACGGCGGAAATTCGTTGAAATTTCCAAATTTATC 180

Dy 121 AAATTTATTTGAGAGTCTGCTGAAACCTACGGCGGAAATTCGTTGAAATTTCCAAATTTATC 180

Qy 181 AAATGTAAAGAAAACCTCTCGACCAATTTCTATGTAAGTGAAGATAAAAAAAGCAAGTGT 240

Dy 181 AAATGTAAAGAAAACCTCTCGACCAATTTCTATGTAAGTGAAGATAAAAAAAGCAAGTGT 240

Qy 241 TCGATAGTCTATTAGATATCGCAACACAGCTCTTTACTTTCCACGTAATAAAAAATTAAT 300

Dy 241 TCGATAGTCTATTAGATATCGCAACACAGCTCTTTACTTTCCACGTAATAAAAAATTAAT 300

Qy 301 CTATCCGAATCAGAAAAACAGTTTAAAAAATAAATGGAATCTATCACTGGGAAAAA 360

Dy 301 CTATCCGAATCAGAAAAACAGTTTAAAAAATAAATGGAATCTATCACTGGGAAAAA 360

Qy 361 TCGGAGAACGCAAAATCAGAAAGGTGGAACTAGTACCCCAAGATTTTCTTAAGATCTTT 420

Dy 361 TCGGAGAACGCAAAATCAGAAAGGTGGAACTAGTACCCCAAGATTTTCTTAAGATCTTT 420

Qy 421 GTTCTTGTCTCCATGTCAGATCATGTTAATGATTTTACATGGTACAAAAATTCGAAAAA 480

Dy 421 GTTCTTGTCTCCATGTCAGATCATGTTAATGATTTTACATGGTACAAAAATTCGAAAAA 480

Db 421 GTTCTTGCTCCATTGCCAGATCATGTTAATGATTTTACATGGTACAAAAATCGAAAAA 480
QY 481 AGCTTAGGTATTAAGCCTGTAAATAGAATATCGGTCCTTCTATTTATTTATTTCTACATTT 540
Db 481 AGCTTAGGTATTAAGCCTGTAAATAGAATATCGGTCCTTCTATTTATTTATTTCTACATTT 540
QY 541 AATCGTAGCGGTATTTAGATATAACGTTAGCGTGTGGTCAATCAGAAAAACAACTAC 600
Db 541 AATCGTAGCGGTATTTAGATATAACGTTAGCGTGTGGTCAATCAGAAAAACAACTAC 600
QY 601 CCATTTGAAGTCGTTGTCAGATGATGGTGAAGAAAAACCTTACTTACCATTGTGCAA 660
Db 601 CCATTTGAAGTCGTTGTCAGATGATGGTGAAGAAAAACCTTACTTACCATTGTGCAA 660
QY 661 AATACGAACAAAACTTGACATAAAGTATGTAAGACAAAAAGATTATGGATATCAATTG 720
Db 661 AATACGAACAAAACTTGACATAAAGTATGTAAGACAAAAAGATTATGGATATCAATTG 720
QY 721 TGTGCAGTCAGAACTTAGGTTTACGTACAGCAAAAGTATGATTGTGTCGATTCTAGAC 780
Db 721 TGTGCAGTCAGAACTTAGGTTTACGTACAGCAAAAGTATGATTGTGTCGATTCTAGAC 780
QY 781 TGCAGATATGSCACCAACAATTTATGGTTTCATCTTATCTTACAGAACTATTAGAAGAC 840
Db 781 TGCAGATATGSCACCAACAATTTATGGTTTCATCTTATCTTACAGAACTATTAGAAGAC 840
QY 841 AATGATATTTGTTTAAATGACCTAGAAAATATGTGGATCTCATATATTACCGCAGAA 900
Db 841 AATGATATTTGTTTAAATGACCTAGAAAATATGTGGATCTCATATATTACCGCAGAA 900
QY 901 CAATTCCTTAACGATCCATATTTAATAGAACTACTACTGTAAGCCGCTACAAAATAAACAAT 960
Db 901 CAATTCCTTAACGATCCATATTTAATAGAACTACTACTGTAAGCCGCTACAAAATAAACAAT 960
QY 961 CCTTCGATTACATCAAAAGGAATATATCTGTGGATTGGAGATTAGAACTATTCAAAAA 1020
Db 961 CCTTCGATTACATCAAAAGGAATATATCTGTGGATTGGAGATTAGAACTATTCAAAAA 1020
QY 1021 ACCGATAATCTACGTCATGATCTCCGTTTCGTTTATTTGTTGCGGTTAATGTTGCA 1080
Db 1021 ACCGATAATCTACGTCATGATCTCCGTTTCGTTTATTTGTTGCGGTTAATGTTGCA 1080
QY 1081 TTTTCTAAAGAAATGGCTAAATAAAGTAGGTTGGTTCGATGAAGAAATTTAATCATTTGGGG 1140
Db 1081 TTTTCTAAAGAAATGGCTAAATAAAGTAGGTTGGTTCGATGAAGAAATTTAATCATTTGGGG 1140
QY 1141 GCGCAAGATGTAGAAATTTGGTTACAGATATTTGCGCAAGGCTGTTTTCAGAGTAAT 1200
Db 1141 GCGCAAGATGTAGAAATTTGGTTACAGATATTTGCGCAAGGCTGTTTTCAGAGTAAT 1200
QY 1201 GACGGCGAATGGCATCCATCAAGAACCCACTGGTAAAGAAAAATGAACAGAACCGGAA 1260
Db 1201 GACGGCGAATGGCATCCATCAAGAACCCACTGGTAAAGAAAAATGAACAGAACCGGAA 1260
QY 1261 GCTGGTAAAGTATTAACGTTTAAATTTGTGAAGAAAGGTACCTTACATCTATAGAAG 1320
Db 1261 GCTGGTAAAGTATTAACGTTTAAATTTGTGAAGAAAGGTACCTTACATCTATAGAAG 1320
QY 1321 CTTTACCATAGAGATTACATATTCATAGATACCTTTAGTTTCTATTTATATCCCC 1380
Db 1321 CTTTACCATAGAGATTACATATTCATAGATACCTTTAGTTTCTATTTATATCCCC 1380
QY 1381 GCTTATAACTGTCAAAATTAATTTCAAGATGTGTAGATAGTGTCTTAACTCAAACTGTT 1440
Db 1381 GCTTATAACTGTCAAAATTAATTTCAAGATGTGTAGATAGTGTCTTAACTCAAACTGTT 1440
QY 1441 GTCGATCTCGAGTTGTTGTTGTAACGATGGTTCAACAGATATATACCTTAGAGTGATC 1500
Db 1441 GTCGATCTCGAGTTGTTGTTGTAACGATGGTTCAACAGATATATACCTTAGAGTGATC 1500
QY 1501 AATAGCTTTATGTTAATAATCTAGGTCAGCATATGTTCTAAACCAAAATGGCGGAATA 1560
Db 1501 AATAGCTTTATGTTAATAATCTAGGTCAGCATATGTTCTAAACCAAAATGGCGGAATA 1560

QY 1561 GCCTCAGCATCAAAATGCAGCCGTTCTTTTGTCTAAAGGTTATTACATTTGGCAGTTAGAT 1620
Db 1561 GCCTCAGCATCAAAATGCAGCCGTTCTTTTGTCTAAAGGTTATTACATTTGGCAGTTAGAT 1620
QY 1621 TCAGATGATTATCTTGAGCCTGATGCACTGTGTTTAAAGAAATTTTAAAGAGAT 1680
Db 1621 TCAGATGATTATCTTGAGCCTGATGCACTGTGTTTAAAGAAATTTTAAAGAGAT 1680
QY 1681 AAAACGCTAGCTTGTGTTTATACCACTTAATAGAAACGTCATCCGGATGTAGCTTAATC 1740
Db 1681 AAAACGCTAGCTTGTGTTTATACCACTTAATAGAAACGTCATCCGGATGTAGCTTAATC 1740
QY 1741 GCTAATGGTTTACAATTTGCCAGAAATTTTTCACGAGAAAACTCACACGGCTATGATGCT 1800
Db 1741 GCTAATGGTTTACAATTTGCCAGAAATTTTTCACGAGAAAACTCACACGGCTATGATGCT 1800
QY 1801 CACATTTTGAATGTTTACGATTAGAGCTTGGCATTTAAGGATGATTTAAGCAAAAT 1860
Db 1801 CACATTTTGAATGTTTACGATTAGAGCTTGGCATTTAAGGATGATTTAAGCAAAAT 1860
QY 1861 ATTGAAACCGCTGGGATTTATGACATGTTCTTAAACTCAGTGAAGTTGGAATAATTTAAA 1920
Db 1861 ATTGAAACCGCTGGGATTTATGACATGTTCTTAAACTCAGTGAAGTTGGAATAATTTAAA 1920
QY 1921 CATCTTAAATAAATCTCTATAACCGGCTATTACATGGTGATAACACATCCATTAAAGAA 1980
Db 1921 CATCTTAAATAAATCTCTATAACCGGCTATTACATGGTGATAACACATCCATTAAAGAA 1980
QY 1981 CTGGGATTTCAAGAAAAACCATTTTGTGTAGTCAATCAGTCATTAATAATAGACAAGC 2040
Db 1981 CTGGGATTTCAAGAAAAACCATTTTGTGTAGTCAATCAGTCATTAATAATAGACAAGC 2040
QY 2041 ATCAATTTATATAATTTATGACAAATTTGATGATTAGATGAAAGTAGAAGTATATCTTC 2100
Db 2041 ATCAATTTATATAATTTATGACAAATTTGATGATTAGATGAAAGTAGAAGTATATCTTC 2100
QY 2101 AATAAAACCGCTGAATATCAAGAAAGAAATGGATATGTTTAAAGATCTTAAACTCATTC 2160
Db 2101 AATAAAACCGCTGAATATCAAGAAAGAAATGGATATGTTTAAAGATCTTAAACTCATTC 2160
QY 2161 AATAAAGATGCAAAATCGCAGTCAGTATTTCTATCCCAATACATTAAGCGCTTAGTG 2220
Db 2161 AATAAAGATGCAAAATCGCAGTCAGTATTTCTATCCCAATACATTAAGCGCTTAGTG 2220
QY 2221 AAAAAACTTAAACAATTTATTTGAATAATAAATAATATTTTCGTTTATTTCTACATGTT 2280
Db 2221 AAAAAACTTAAACAATTTATTTGAATAATAAATAATATTTTCGTTTATTTCTACATGTT 2280
QY 2281 GATAGAATCATCTTACACGAGACATCAAAAAAGAAATTTGGCTTCTATCATTAAGCAC 2340
Db 2281 GATAGAATCATCTTACACGAGACATCAAAAAAGAAATTTGGCTTCTATCATTAAGCAC 2340
QY 2341 CAAGTGAATATTTTACTAAATAAGTACATCTCATATTACAGAGTAATAGACTAATAAAA 2400
Db 2341 CAAGTGAATATTTTACTAAATAAGTACATCTCATATTACAGAGTAATAGACTAATAAAA 2400
QY 2401 ACTGAGGCACATTTAAGTAAATTTAATAAATAAAGTCAAGTAAATCTAAATTTGTGAATAC 2460
Db 2401 ACTGAGGCACATTTAAGTAAATTTAATAAATAAAGTCAAGTAAATCTAAATTTGTGAATAC 2460
QY 2461 ATCAATTTTGTATCATGACAGCCTATTCGTTTAAATAAGCAGCTATGCTTTATATGAAA 2520
Db 2461 ATCAATTTTGTATCATGACAGCCTATTCGTTTAAATAAGCAGCTATGCTTTATATGAAA 2520
QY 2521 AAATATGATGTGCGCATGAATTTCTCAGCATTTAACACATGATTGGATCGAGAAAAATCAAT 2580
Db 2521 AAATATGATGTGCGCATGAATTTCTCAGCATTTAACACATGATTGGATCGAGAAAAATCAAT 2580
QY 2581 GCGCATCCACCATTTTAAAGAGCTGATTTAAACCTATTTTAAATGACAAAGTATAGAAGT 2640
Db 2581 GCGCATCCACCATTTTAAAGAGCTGATTTAAACCTATTTTAAATGACAAAGTATAGAAGT 2640

Db	901	TTCTTGTGAATCATTACCAAGAAGTGAARACCAATAATAGTTGTCGCGAAAGGGGAAGCA	960
Qy	982	AATATATACGTTGGATTTGAGATTTAGAACAATTTCAAAAAACCGATATATCTACGCTCTATGT	1041
Db	961	ACAGTTTCTCTGGATTGGCGCTTAGAACAAATTCGAAAAAACAAGAAATCTCGCCTTATCC	1020
Qy	1042	GATTCCTCGTTTCGTTATTTTCTTCGGGTAATGTGCGATTTTCTAAGAATAGCCTAAAT	1101
Db	1021	GATTCGCGTTTCGTTTTCGCGGGGTAAATGTTCCTTCGCTAAAAAATGCGCTAAAT	1080
Qy	1102	AAAGTAGGTTGGTTCGATGAAGAAATTTAATCATTTGGGGGGCGGAAGATGTAGAAATTTGGT	1161
Db	1081	AAATCCGGTTTCTTTGATGAGGAATTTAATCACTGGGCTGGAAGAATGTGGAATTTGGA	1140
Qy	1162	TACAGATTTATTTGCCAAGGCTGTTTTTTCAGAGTAATTTGACGCGCGGAATGGCCATCCAT	1221
Db	1141	TATCGCTTTATTCGGTTTACGGTAGTTCTTTTAAACTATTTGATGCGATTTATGGCCTACCAT	1200
Qy	1222	CAAGAACCACCTGSTAAAGAAAATGAACAGAAGCGGAAGCTGGTAAAAAGTATTACGCTT	1281
Db	1201	CAAGGCCACCAAGTTAAACAAAATGAACCCGATCGTGAAGCGGAAAAAATATTACGCTC	1260
Qy	1282	AAATTTGTGAAGAAAAGTACCTTTACATCTATAGAAGCTTTTACCAATAGAGATTCA	1341
Db	1261	GATATTATGAGAAAAGTCCCTTATATCTATAGAAAACCTTTTACCANATAGAAAGATCG	1320
Qy	1342	CATATTCATAGAATACCTTTAGTTTCTATTTTATATCCCGCTTATACTGTGCAAAATTAAT	1401
Db	1321	CATATCAATAGATACCTTTAGTTTCAATTTATATCCAGCTTATAACTGTGCAAACTAT	1380
Qy	1402	ATTCAAGATGCTAGATAGTGCCTTTAATCAAACTGTTGTCGATCTCGAGGTTTGTGATT	1461
Db	1381	ATTCAAGCTTTCGCTAGATAGTGCACTGAATCAGACTGTTGTGATCTCGAGGTTTGTATT	1440
Qy	1462	TGTAAAGATGGTTTCAACAGATAATACCTTTAGAAGTGATCAATAAGCTTTTATGTTAAATTAAT	1521
Db	1441	TGTAAAGATGGTTTCAACAGATAATACCTTTAGAAGTGATCAATAAGCTTTTATGTTAAATTAAT	1500
Qy	1522	CCTAGGCTAGCATCATGTCTAAACCAATGCGCGAATAGCCTCAGCATCAAAATGCAACC	1581
Db	1501	CCTAGGCTAGCATCATGTCTAAACCAATGCGCGAATAGCCTCAGCATCAAAATGCAACC	1560
Qy	1582	GTTTCTTTTTCGCTAAAGTTATTTACATTTGGCGAGTTTAGATTTCAGATGATTATCTTGAGCCT	1641
Db	1561	GTTTCTTTTTCGCTAAAGTTATTTACATTTGGCGAGTTTAGATTTCAGATGATTATCTTGAGCCT	1620
Qy	1642	GATCAGCTTGAACGTGTTTAAAGAAATTTTAAAGATAAAACCGCTAGCTTTGTGTTTAT	1701
Db	1621	GATCAGCTTGAACGTGTTTAAAGAAATTTTAAAGATAAAACCGCTAGCTTTGTGTTTAT	1680
Qy	1702	ACCCTAATAGAACGTCATCCGGATCGTAGCTTAATCGCTTAATGTTTACAATTTGGCCA	1761
Db	1681	ACCCTAATAGAACGTCATCCGGATCGTAGCTTAATCGCTTAATGTTTACAATTTGGCCA	1740
Qy	1762	GAATTTTTCAGGAAAAAAGCTCACAACGGGTATGTTGCTACCACTTTTAGAATGTTTACG	1821
Db	1741	GAATTTTTCAGGAAAAAAGCTCACAACGGGTATGTTGCTACCACTTTTAGAATGTTTACG	1800
Qy	1822	ATTAGAGCTTGGCATTTTAAACGGATGATTTACGGAATATTTGAAACGCGCTGGATTAT	1881
Db	1801	ATTAGAGCTTGGCATTTTAACTGATGATTCATGAAATAATTTGAAATTTGCCGTAGCAT	1860
Qy	1882	GACATGTTCTTAAACTCAGTGAAGTTGAAAAATTTAAACATCTTTAAATAAATCTGCTAT	1941
Db	1861	GACATGTTCTTAAACTCAGTGAAGTTGAAAAATTTAAACATCTTTAAATAAATCTGCTAT	1920
Qy	1942	AACCGGTATTACATGGTGATACACATCCATTAGAAGAACTCGGCTTCAAGAGAAAAC	2001
Db	1921	AACCGGTATTACATGGTGATACACATCAATTAAGAAACTTTGGCATTTCAAGAGAAAAC	1980
Qy	2002	CATTTGTTGTAGTCAATCAGTCAATTAATAGAACGGCATCAATTTATTAATTTATGAC	2061

1981	DB	CA	TTTTGTGTAGTCAATTAATAATAGCAAGGCATTAACCTTATTATTATTATTATATGAC	2040
2062	QY	AA	ATTTTGATGATTTAGATGAAAGTAGAAGTATATCTTCAATAAAACCGCTGAATATCAA	2121
2041	DB	GA	ATTTTGATGATTTAGATGAAAGTAGAAGTATATTTTCAATAAAACCGCTGAATATCAA	2100
2122	QY	GA	AGAAATGGATATGTTAAAGATCTTAAACHTCATTCAAAAATAAGATGCCAAATCGCA	2181
2101	DB	GA	AGAGATGGATATCTTAAAGATATTAAATCATCCAGAATAAGATGCCAAATCGCA	2160
2182	QY	GT	CAGTATTTTCTATCCCAATACATTAACGGCTTAGTGAAAAAACTAAACAATATTAT	2241
2161	DB	GT	CAGTATTTTATCCCAATACATTAACGGCTTAGTGAAAAAACTAAACAATATTAT	2220
2242	QY	GA	ATATAATAAAAAATATATTCGTATTATTCTACATGTTGATAAGAATCATCTTACACCA	2301
2221	DB	GA	ATATAATAAAAAATATATTCGTATTGTTCTACATGTTGATAAGAATCATCTTACACCA	2280
2302	QY	GAC	ATCAAAAAAGAAATATTGGCTTCTATCATATAAGCACCAAGTGAATATTTTACTAAAT	2361
2281	DB	GAT	ATCAAAAAAGAAATACTAGCCTCTCATATAACATCAAGTGAATATTTTACTAAAT	2340
2362	QY	AAT	GACATCTCATATTACACGAGTATAGACHTAATAAAACCTGAGGCACATTTTAAAGTAAT	2421
2341	DB	AAT	GATCTCATATTACACGAGTATAGATTAAATAAAACCTGAGGCATTTTAAAGTAAT	2400
2422	QY	ATT	AATAAATTAAGTCAGTTAAATCTGAAATGTTGAATACATCATTTTTTGATAATCATGAC	2481
2401	DB	ATT	AATAAATTAAGTCAGTTAAATCTGAAATGTTGAATACATCATTTTTTGATAATCATGAC	2460
2482	QY	AG	CTATTCGTTAAAAATGACAGCTATGCTTATATGAAAAAATATGATGTCGCATGAAT	2541
2461	DB	AG	CTATTCGTTAAAAATGACAGCTATGCTTATATGAAAAAATATGATGTCGCATGAAT	2520
2542	QY	TT	CTCAGCATTAACACATGATTCGGATCGAGAAATCAATCGCATCCACCATTTAAAAAG	2601
2521	DB	TT	CTCAGCATTAACACATGATTCGGATCGAGAAATCAATCGCATCCACCATTTAAAAAG	2580
2602	QY	CT	GATTTAAACCTATTTTATGACAATGACTTTAAGAAGTATGAATGTGAAAGGGGCATCA	2661
2581	DB	CT	GATTTAAACCTATTTTATGACAATGACTTTAAGAAGTATGAATGTGAAAGGGGCATCA	2640
2662	QY	CA	AGGTATGTTATGAAGTATGCGCTACCGCATGAGCTTCTGACGATTATTAAAGRAGTC	2721
2641	DB	CA	AGGTATGTTTATGACGTATGCGCTAGCGCATGAGCTTCTGACGATTATTAAAGRAGTC	2700
2722	QY	AT	CACATCTCTGCCAATCAATTGTAGTGTGCCAGAATATAACACTGAGGATATTTGGTTC	2781
2701	DB	AT	CACATCTTTGCCAGTCAATTGTAGTGTGCCAGAATATAACACTGAGGATATTTGGTTC	2760
2782	QY	CA	ATTTTGCACTTTTAACTTTAGAAAAAGAAACCGGCATGTTTATAAAACATCGACC	2841
2761	DB	CA	ATTTTGCACTTTTAACTTTAGAAAAAGAAACCGGCATGTTTATAAAACATCGACC	2820
2842	QY	CT	GACTTATATGCGTTGGGAACGAAATTTACATGGACAATGNACAAATTCAAAGTGCA	2901
2821	DB	CT	GACTTATATGCGTTGGGAACGAAATTTACATGGACAATGNACAAATTTGAAAGTGCA	2880
2902	QY	AAAA	AGCGCAAAATATCCCGTTTAAACAAGTTCATTATTAAATAGTATAACCGCTATAA	2958
2881	DB	AAA	AGGAGGAGAAATATACCTGTTAACAGTTCATTATTAAATAGTATAACCTCTATAA	2937

RESULT 5
AAZ35589
ID AAZ35589 standard: DNA: 2937 BP.

p multocida hyaluronate synthase (PMHAS) nucleotide sequence.

HW hyaluronate synthase; PmHAS; hyaluronan; hyaluronic acid; HA; cosmetic;
KW drug delivery; angiogenesis; wound healing; capsule synthesis;
KW fowl cholera; shipping fever; ss.

OS Pasturella multocida.

XX Key Location/Qualifiers
XX CDS 19..2937
XX FT /*tag= a
XX FT /product= PmHAS
XX FT /note= "Hyaluronate synthase"

XX W09951265-A1.

XX PD 14-OCT-1999.

XX PF 01-APR-1999; 99WO-US07289.

XX PR 02-APR-1998; 98US-0080414.

XX PR 26-OCT-1998; 98US-0178851.

XX PA (OKLA) UNIV OKLAHOMA.

XX PI Deangelis P;

XX WPI; 2000-013032/01.

XX DR P-PSDB; AAY43099.

XX New isolated hyaluronate synthase nucleic acids, used for the
PT production of hyaluronic acid, for developing antibiotics and vaccines
PT and for diagnostic applications -

XX Claim 3; Page 113-114; 121pp; English.

XX This is the Pasturella multocida hyaluronate synthase (PmHAS) nucleic
CC acid sequence. Hyaluronic acid (HA) or hyaluronan, is a polysaccharide
CC that serves both structural and recognition roles in higher animals.
CC Bacteria produce extracellular capsules of HA which mimic their host HA
CC and aid escape from a host immune response. The invention includes a
CC vector containing the PmHAS nucleotide sequence which can be used to
CC express PmHAS in a foreign host. The HS nucleic acids can be used for
CC production of HA. Also, specific changes to the HS coding sequence can
CC result in the production of HA having a modified size distribution or
CC structural configuration and functional properties. The HA products can
CC be used in e.g. drug delivery, angiogenesis and wound healing,
CC stabilisation of recombinant proteins and in cosmetics. The HS nucleic
CC acids can also be used to develop agents to block capsule synthesis by
CC pathogens and act as antibiotics. The avirulent P. multocida strains can
CC be used as vaccines for fowl cholera or shipping fever.

XX SQ Sequence 2937 BP; 1077 A; 482 C; 509 G; 869 T; 0 other;

Query Match 76.68; Score 2283.4; DB 21; Length 2937;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 2540; Conservative 0; Mismatches 376; Indels 21; Gaps 1;

Qy 43 ATTTTAAAGGAAAGAAATCAATATATACAGCAATATATACAGCAATATATACAGCAAT 102

Db 1 ATTTTAAAGGACGAAATCAATATATACAGCAATATATACAGCAATATATACAGCAAT 60

Qy 103 GACTATGAATTAGCACTCAAAATTTATTTGAGAGTCTGCTGAAACCTACGGCGGAAAAATC 162

Db 61 GACTATCAATTAGCACTCAAAATTTATTTGAAAGTCGGCGGAAATCTATGGCGGAAATTT 120

Qy 163 GTTGAATTCGAATATCAATATGTAAGAAAAAATCTCGACCAATTC----- 209

Db 121 GTTGAATTTCAATATACCAATATGCAAGAAAAATCTCTCAGCACATCTCTGTTTAATTC 180

Qy 210 -----TTATGTAAAGTGAAGATAAAAAACACAGTGTTCGATAGCTCATAGATATC 261

Db 181 GCACATCTTCTGTAATTAAGAGAAAAAGTCAATGTTTGGGATAGTCGTTAGATATT 240

Qy 262 GCAACACAGCTCTTACTTTTCCAACGTAATAAAAAATTAACCTCTATCCGAATCAGAAAAAAC 321

Db 241 GCAACACAACTGTTACTTTTCCAACGTAATAAAAAATTAGTACTTCTGACTCGGAAAAAAC 300

Qy 322 AGTTTAAAAAATAAATGGAATCTATCACTCGGAAAAAATCGAGAACGACGAAATCAGA 381

Db 301 ACGTTAAAAAATAAATGGAATTTGCTCACTGAGAAAAATCTGAAATTCGAGGTAAGA 360

Qy 382 AAGGTGGAACACTAGTACCCAAAGATTTTCCATAAGATCTTGTCTTCTGCTCCATTTGCCAGAT 441

Db 361 GGGGTGCGCCTTGTACCAAAAGATTTTCCCAAAGATCTGTGTTTACGCGCTTACCTGAT 420

Qy 442 CATGTTAATGATTTTACATGTTACAAAAATCGAAAAAAGCTTAGGTATAAAGCCTGTA 501

Db 421 CATGTTAATGATTTTACATGTTACAAAAAGCGAAAAAGAACTTGGCATAAAAACCTGAA 480

Qy 502 ATAAGAATATCGGTCTTCTTCTATTTATTTCTACATTTTAACTCGTAGCGGTATTTAGAT 561

Db 481 CATCAACATGTTGGTCTTTTCTATTTATTCGTTACAACATTCATCGACGACCAATTTTATCG 540

Qy 562 ATAAGCTTAGCCTGTTTGGTCAATCAGAAAAAACAATACCCATTTTGAAGTCTGTTGTGCA 621

Db 541 ATTACATTAGCCTGTTTGTAAACCAAAAAACACATTTACCGTTCGAAGTTATCGTGACA 600

Qy 622 GATGATGTTAGTAAAGAAAACTTACTTTACCATTTGTGCAAAAAATACGAAACAACTTTGAC 681

Db 601 GATGATGTTAGTCAGGAAGATCTATCACCGATCATTCACGCAATATCGCAATATGAAATAAATGGAT 660

Qy 682 ATAAGATGTTAAGACAAAAAGATTTATGATATGATATCAATTTGTCGACTGCGACTCAGAACTTAGGT 741

Db 661 ATTCCTGACTGACGAAAAAAGATTAACGGTTTTCAGCCAGTGCCTCGGAAATATGGGA 720

Qy 742 TTACGTACAGCAAAAGTATGATTTTGTCTCGATTTCTAGCTGCGATATGCGACCAACAACAA 801

Db 721 TTACGCTTAGCAAAATATGACTTTATTGGCTTACTCGACTGTGATATGGCCCAATATCCA 780

Qy 802 TTATGGTTCATTTCTTATCTTACAGAACTTATTAGAGACAATGATATTTGTTTAAATGGA 861

Db 781 TTATGGTTCATTTCTTATCTTGCAGAGCTATTAGAAAGATGATGATTAAACAATCATTTGGT 840

Qy 862 CCTAGAAATATGTCGATCTCATATATATACGCGAAGAAATTCCTTTACGATCCATAT 921

Db 841 CCAAGAAAAATACGATGATACAAACATATGACCCAAAAGACTTTTAAATAACGGCAGT 900

Qy 922 TTAATAGATCACTACTACCTGAAACCGCTACAAATAACAATCTTTCGATTTACATCAAAAGGA 981

Db 901 TTGCTTGAATCATTACCAGAACTGAAACCAATAATAGTTGTTCCCGCAAAAGGGAGGA 960

Qy 982 AATATATCGTTGGATGGAGATTAGAACATTTCAAAAAAACCGATTAATCTACCTCTATGT 1041

Db 961 ACAGTTTCTCTGGATTGGCGCTTAGAACAAATTCGAAAAAACAGAAATCTCCGCTTATCC 1020

Qy 1042 GATTCCTCGTTTCGTTATTTTGTTCGGGTAATGTTGCAATTTTCTAAAGATGCTTAAAT 1101

Db 1021 GATTCGCTTTCGTTTTCGCGGGTAAATGTTGCTTTCGCTTAAAAAATGGCTAAAT 1080

Qy 1102 AAAGTAGGTGTTGTCATGAAGAAATTAATCATTTGGGGGGCGGAAGATGTAGAAATTTGGT 1161

Db 1081 AAATCGGTTTCTTGTATGAGNAATTAATCACTGGGGTGGAGAGATGTGGAATTTGGA 1140

Qy 1162 TCAGATTAATTTGCCAAAGCTGTTTTTTCAGAGTAATTTGACGCGGGAATGGCCATCCAT 1221

Db 1141 TATCGCTTATTCGTTTACGCTAGTTTCTTTTAAACTATTGATGGCATTTATGGCTACCAT 1200

Qy 1222 CAAGAACACCTTGGTAAAGAAATGAACAGACGCGAGCTGTAAGAGTATTACGCTT 1281

Db 1201 CAAGACCCACAGGTAAAGAAATGAACCGATCGTGAAGCGGGAAGAAATATTACGCTC 1260

Qy 1282 AAAATTTGAAAGAAAGGTACTTACATCTATAGAAAGCTTTTACCAATAGAAAGATTCA 1341

Db 1261 GATATTATGAGAAAAAGGTCCCTTATATCTATAGAAAACTTTTACCAATAGAAAGATTG 1320

Qy 1342 CATATTATAGAAATACCTTTTAGTTTCTTATTTATATCCCGCTTATAACTGTGCAAAATAT 1401

Db 1321 CATATCAATAGAGTACCTTTAGTTTCAATTTATATATCCAGCTTATAAAGTGTGCAAACTAT 1380
QY 1402 ATTCAAGATGTAGATAGTCTCTTAATCAAACTGTTGTGCACTCGAGCTTGTGTAAT 1461
Db 1381 ATTCAAGCTTGGGTAGATAGTGCATGAATCAGACTGTTGTTGATCTCGAGGTTGTGTAAT 1440
QY 1462 TCTAACGATGTTCAACAGATTAATACCTTTAGAAAGTGATCAATAAGCTTTATGGTAATAAT 1521
Db 1441 TGTACAGATGGTTCAACAGATTAATACCTTTAGAAAGTGATCAATAAGCTTTATGGTAATAAT 1500
QY 1522 CCTAGGTTACGATCATGCTTTAAACCAATAGCGGGAATAGCCTCAGCATCAAAATGAGCC 1581
Db 1501 CCTAGGTTACGATCATGCTTTAAACCAATAGCGGGAATAGCCTCAGCATCAAAATGAGCC 1560
QY 1582 GTTCTTTTGGCTTAAGGTTATTTACATTTGGCAGTTAGATTCAGATGATTTATCTTGAGCCT 1641
Db 1561 GTTCTTTTGGCTTAAGGTTATTTACATTTGGCAGTTAGATTCAGATGATTTATCTTGAGCCT 1620
QY 1642 GATGAGTTGAACTGTGTTTAAAGAAATTTTAAAGATAAAACGCTAGCTTGTGTTTAT 1701
Db 1621 GATGAGTTGAACTGTGTTTAAAGAAATTTTAAAGATAAAACGCTAGCTTGTGTTTAT 1680
QY 1702 ACCACTAATAGAACTCAATCGGATGTAGCTTAATCGCTAAAGTTACAAATTTGGCCA 1761
Db 1681 ACCACTAATAGAACTCAATCGGATGTAGCTTAATCGCTAAAGTTACAAATTTGGCCA 1740
QY 1762 GAATTTTCAGGAAAGAACTCAACGCGCTATGATGCTCACCATTTTAGAAATGTTTACG 1821
Db 1741 GAATTTTCAGGAAAGAACTCAACGCGCTATGATGCTCACCATTTTAGAAATGTTTACG 1800
QY 1822 ATTAGAGCTTGGCATTTAACGGATGATTTAACGAAATATTGAAACGCGCTGGATAT 1881
Db 1801 ATTAGAGCTTGGCATTTAACGGATGATTTAACGAAATATTGAAACGCGCTGGATAT 1860
QY 1882 GACATGTTCTTAACTCAGTGAAGTTGGAATTTTAAACATCTTAAATAAATCTGCTAT 1941
Db 1861 GACATGTTCTTAACTCAGTGAAGTTGGAATTTTAAACATCTTAAATAAATCTGCTAT 1920
QY 1942 AACCGGTATTACATGGTGATACACATCCATTAAAGAACTCGGCATTTCAAAAGAAAAAC 2001
Db 1921 AACCGGTATTACATGGTGATACACATCCATTAAAGAACTCGGCATTTCAAAAGAAAAAC 1980
QY 2002 CATTTTGTGTAGTCAATCAGTCAATTAATAGACAGGCAATCAATTAATAATATTAGAC 2061
Db 1981 CATTTTGTGTAGTCAATCAGTCAATTAATAGACAGGCAATCAATTAATAATATTAGAC 2040
QY 2062 AAATTTGATGATTTAGATGAAGTGAAGTATCTTCAATAAAACCGCTGAATATCAA 2121
Db 2041 GAATTTGATGATTTAGATGAAGTGAAGTATCTTCAATAAAACCGCTGAATATCAA 2100
QY 2122 GAAGAAATGGATATGTTAAAGATCTTAAACTCATTCAAAATAAAGATGCCAAATCGCA 2181
Db 2101 GAAGAGATTGATATCTTAAAGATATTAATAATCATCCAGAATAAGATGCCAAATCGCA 2160
QY 2182 GTGAGTATTTTATCCCAATACATTAAGCGGCTTAGTGAAGAACTAAACATATTTATT 2241
Db 2161 GTGAGTATTTTATCCCAATACATTAAGCGGCTTAGTGAAGAACTAAACATATTTATT 2220
QY 2242 GAATAATAAATAATATTTCTGTTTATTTCTACATCTTGAAGAATCATCTTACACCA 2301
Db 2221 GAATAATAAATAATATTTCTGTTTATTTCTACATCTTGAAGAATCATCTTACACCA 2280
QY 2302 GACATCAAAAAGAAATATTGGCTTTCTATCATAGAACCAAGTGAATATTTTACTAAAT 2361
Db 2281 GATATCAAAAAGAAATATTAGCCTTCTATCATAAACATCAAGTGAATATTTTACTAAAT 2340
QY 2362 AATGACATCTCATATTTACAGAGTAATAGACTAATAAAGCTGAGGCACATTTAAGTAAT 2421
Db 2341 AATGATATCTCATATTTACAGAGTAATAGACTAATAAAGCTGAGGCACATTTAAGTAAT 2400
QY 2422 ATTAATAAATTAAGTCAAGTAAATCTAAATTTGTAATACATCATTTTGTGATAATCATGAC 2481
Db 2401 ATTAATAAATTAAGTCAAGTAAATCTAAATTTGTAATACATCATTTTGTGATAATCATGAC 2460

QY 2482 AGCCTATTCGTTAAAAATGACAGCTATGCTTATATGAAAAAATATATGATGCGCATGAAT 2541
Db 2461 AGCCTATTCGTTAAAAATGACAGCTATGCTTATATGAAAAAATATGATGCGCATGAAT 2520
QY 2542 TTCTCAGCATTAACACATGATTTGGATCGAGAAATCAATGCGCATCCACCATTTAAAAAG 2601
Db 2521 TTCTCAGCATTAACACATGATTTGGATCGAGAAATCAATGCGCATCCACCATTTAAAAAG 2580
QY 2602 CTGATTTAAACCTATTTTAAATGACAATGACTTAAAGATGATGAATGTGAAGGGGCATCA 2661
Db 2581 CTCATTTAAACCTATTTTAAATGACAATGACTTAAAGATGATGAATGTGAAGGGGCATCA 2640
QY 2662 CAAGGTATGTTTATGAAGTATGCGTACCGCATGAGCTTCTGACGATTTATTAAGAAAGTC 2721
Db 2641 CAAGGTATGTTTATGAAGTATGCGTACCGCATGAGCTTCTGACGATTTATTAAGAAAGTC 2700
QY 2722 ATCAGATCTGCGCAATCAATTTGATGTTGCGCAGATATACACTGAGGATATTTGGTTC 2781
Db 2701 ATCAGATCTGCGCAGTCAATTTGATGTTGCGCAGATATACACTGAGGATATTTGGTTC 2760
QY 2782 CAATTTGCACTTTTAACTCTTAGAAAAAGAAACCGGCCATGATTTTAAATAAAACATCGACC 2841
Db 2761 CAATTTGCACTTTTAACTCTTAGAAAAAGAAACCGGCCATGATTTTAAATAAAACATCGACC 2820
QY 2842 CTGACTTATATGCTTGGGAAACGAAATTAACAATGAGCAAAATGACAAATTTCAAAGTGCA 2901
Db 2821 CTGACTTATATGCTTGGGAAACGAAATTTACAATGAGCAAAATGACAAATTTGAAAGTGCA 2880
QY 2902 AAAAAAGCGCAAAATATCCCGTTTAAACAAGTTCAATTTAATATATATACGCTATAA 2958
Db 2881 AAAAAAGCGCAAAATATACCTGTTAACAGTTTCATTTAATATATAGTATACCTATAA 2937

RESULT 6
AAA27448
ID AAA27448 standard; cdna; 2112 BP.
XX
AC AAA27448;
XX AC
XX 17-AUG-2000 (first entry)
DE P.mutocida hyaluronic acid synthase-D coding sequence.
XX
XX Hyaluronic acid synthase-D; HAS-D; enzyme; ophthalmic surgery;
KW cataract; arthritis; ulcer; tissue abrasion; viscoelastic replacement;
KW hyaluronic acid production; bioadhesive; ss.
XX
XX Pasteurella multocida.
XX
XX Key Location/Qualifiers
CDS 1..2112
FT /*tag= a
FT /product= "PmHAS-D"
FT /trans_except= (pos:1348..1353,aa:Asn)
XX
PN WO200027437-A2.
XX
XX 18-MAY-2000.
XX
XX 10-NOV-1999; 99WO-US26501.
XX
PR 11-NOV-1998; 98US-0107929.
PR 01-APR-1999; 99US-0283402.
XX
XX (OKLA) UNIV OKLAHOMA STATE.
XX
XX DeAngelis PL;
XX WPI; 2000-376319/32.
DR P-PSDB; AAY96212.
XX
XX Novel method for the enzymatic transfer of sugar molecules to an

FT CDS /note= "encodes AAY54065"
 3390..4121
 /*tag= b
 /product= "EPS2"
 /note= "encodes AAY54066"
 4130..4822
 /*tag= c
 /product= "EPS3"
 /note= "encodes AAY54067"
 4832..5530
 /*tag= d
 /product= "EPS4"
 /note= "encodes AAY54068"
 5629..6996
 /*tag= e
 /product= "EPS5"
 /note= "encodes AAY54069"
 8633..10681
 /*tag= f
 /product= "EPS6"
 /note= "encodes AAY54070: contains 1 stop codon at
 nucleotides 9410-9412"
 10703..11671
 /*tag= g
 /product= "EPS7"
 /note= "encodes AAY54071"
 12838..13788
 /*tag= h
 /product= "EPS8"
 /note= "encodes AAY54072"
 14138..15553
 /*tag= i
 /product= "EPS9"
 /note= "encodes AAY54073"
 16919..18016
 /*tag= j
 /product= "EPS10"
 /note= "encodes AAY54074"
 XX
 PN W09962316-A2.
 XX
 PD 09-DEC-1999.
 XX
 PF 22-APR-1999; 99WO-EP02841.
 XX
 PR 22-APR-1998; 98EP-0201310.
 PR 22-APR-1998; 98EP-0201311.
 PR 22-APR-1998; 98EP-0201312.
 XX
 PA (NEST) SOC PROD NESTLE SA.
 XX
 PI Stingele F, Germond JE, Lamothe G;
 XX
 DR WP; 2000-097267/08.
 DR P-PSDB; AAY54065, AAY54066, AAY54067, AAY54068, AAY54069, AAY54070,
 DR AAY54071, AAY54072, AAY54073, AAY54074.
 XX
 XX New recombinant enzymes for synthesis of exopolysaccharides,
 PT particularly in lactic acid bacteria, for improving properties of
 PT fermented milk products
 XX
 XX Claim 6; Page 77-96; 162pp; French.
 XX
 CC The present sequence encodes enzymes involved in the biosynthesis of
 CC exopolysaccharides (EPS). These enzymes are designated EPS1-EPS10. and
 CC are encoded by open reading frames eps1-eps10. The enzymes are isolated
 CC from *Streptococcus thermophilus* strain Sfi39. The proteins are used
 CC in a method for the synthesis of EPS, which includes at least one step
 CC of forming a bond (alpha or beta-isomer) between C-1 (carrying the
 CC reducing aldehyde function, of an activated D-galactose pyranose), and
 CC a phosphate on a lipophilic or proteinaceous primer. Stepwise synthesis
 CC of EPS occurs with, in each step, addition of a new sugar unit, through
 CC its hemi-acetyl function, to an alcoholic hydroxyl of a second sugar

CC unit, present at the end of a chain of sugar residues bonded to the
 CC primer. EPS1 to EPS4 are used to elongate the EPS chain, and to regulate
 CC EPS production, EPS5 creates new bonds between saccharides, EPS6 is used
 CC in the biosynthesis of EPS, EPS7 and EPS8 are beta-glycosyltransferases,
 CC EPS9 is a transporter of repetitive units, and EPS10 catalyses the
 CC conversion of a pyranose form of a beta-D-galactose to the furanose
 CC form. The EPS enzyme are used to improve properties of foods,
 CC particularly fermented milk products such as yoghurt and cheese,
 CC e.g. their organoleptic properties and flavour stability.
 XX
 SQ Sequence 19966 BP; 6405 A; 3134 C; 3732 G; 6694 T; 1 other;
 Query Match 2.7%; Score 80; DB 21; Length 19966;
 Best Local Similarity 52.0%; Pred. No. 4.1e-06;
 Matches 179; Conservative 0; Mismatches 165; Indels 0; Gaps 0;
 OY 1307 ACATCTATAGAAAGCTTTTACCAATAGAGATTACACATATTCATAGAAATACCTTTAGCTTT 1366
 DB 12794 ACTGTTAGAAATATGATTTTATATTAATTTAGGAGTAGAATAAGAGATCAATCCATTATAT 12853
 OY 1367 CTATTTATATCCCGCTTTTACCAATAGAGATTACCAATATATATTCAAAGATGTGTAGATGCTC 1426
 DB 12854 CAATTATTTGTTCCACATATAATGTTGAAATATATATTAGACATGATTGAATCAATCT 12913
 OY 1427 TTAATCAACTGTTCGATCTCGAGTTTGTATTTCGTAACGATGGTTCACACATATA 1486
 DB 12914 TAGCTCAACATATCGCAATATTGAAGTCATTATAGTATGATGGTAGACAGATCAGT 12973
 OY 1487 CCTAGAGTGATCAATAAGCTTTTATGTTAATATCTAGGTTACGATCATGTCTAAAC 1546
 DB 12974 CGTAGCAGTAATTTCCGATTTAATCTGTAGTCATCATATATTAAGTATCAACCAA 13033
 OY 1547 CAAATGCGGAATAGCCTCAGCATCAATAGCAGCGTTCTTTTGGCTAAAGGTTATTACA 1606
 DB 13034 AAAACCAAGGATTATATAGTAGTCTGAAACACTGTTATGATCGCGCAACTGGTAAATATA 13093
 OY 1607 TTGGCGAGTTAGATTTCAGATGATTATCTTGAGCTCGATGTCAGTT 1650
 DB 13094 TAGCTTTGTAGATGCGATGACAAATAATTAAGCCAGACTTTGTT 13137
 RESULT 11
 AAZ30355
 ID AAZ30355 standard; DNA; 19966 BP.
 XX
 AC AAZ30355;
 XX
 DT 11-FEB-2000 (first entry)
 XX
 DE Nucleotide sequence of the eps operon of *S. thermophilus* Sfi39.
 XX
 KW eps operon; *Streptococcus thermophilus* Sfi39; enzyme; eps1; eps2;
 KW eps3; eps4; eps5; eps6; eps7; eps8; eps9; eps10;
 KW exopolysaccharide biosynthesis; EPS; intersugar bond; antitumour;
 KW probiotic; foodstuff; organoleptic quality; flavour;
 KW lactic acid bacteria; acidified milk product; yoghurt; cheese; ss.
 XX
 OS *Streptococcus thermophilus*.
 XX
 XX Key Location/Qualifiers
 XX 2703..3389
 XX CDS /*tag= a
 XX /product= "eps1"
 XX /note= "encodes AAY43767"
 XX 3390..4121
 XX CDS /*tag= b
 XX /product= "eps2"
 XX /note= "encodes AAY43768"
 XX 4130..4822
 XX CDS /*tag= c
 XX /product= "eps3"
 XX /note= "encodes AAY43769"
 XX 4832..5530


```
FT FT /*tag= d
FT FT /product= "eps4"
FT FT /note= "encodes AAY43770"
FT CDS 5629..6996
FT FT /*tag= e
FT FT /product= "eps5"
FT FT /note= "encodes AAY43771"
FT CDS 8633..10681
FT FT /*tag= f
FT FT /product= "eps6; this CDS, minus the termination
FT FT . codon, is specifically claimed (claim 6)"
FT FT /note= "encodes AAY43772; contains a stop codon at
FT FT nucleotides 9410-9412"
FT CDS 10703..11671
FT FT /*tag= g
FT FT /product= "eps7"
FT FT /note= "encodes AAY43773; this CDS, minus the
FT FT termination codon, is specifically claimed
FT FT (claim 6)"
FT CDS 12838..13788
FT FT /*tag= h
FT FT /product= "eps8"
FT FT /note= "encodes AAY43774; this CDS, minus the
FT FT termination codon, is specifically claimed
FT FT (claim 6)"
FT CDS 14138..15553
FT FT /*tag= i
FT FT /product= "eps9"
FT FT /note= "encodes AAY43775; this CDS is specifically
FT FT claimed (claim 6)"
FT CDS 16919..18016
FT FT /*tag= j
FT FT /product= "eps10"
FT FT /note= "encodes AAY43776; this CDS, minus the
FT FT termination codon, is specifically claimed
FT FT (claim 6)"
FT FT
FT XX W09954475-A2.
FT XX
FT XX 28-OCT-1999.
FT XX
FT XX 22-APR-1999; 99WO-EP03011.
FT XX
FT XX 22-APR-1998; 98EP-0201310.
FT PR 22-APR-1998; 98EP-0201311.
FT PR 22-APR-1998; 98EP-0201312.
FT XX
FT PA (NEST ) SOC PROD NESTLE SA.
FT XX
FT XX Stingele F, Germond JE, Lamothe G;
FT XX
FT XX WPI: 2000-013255/01.
FT DR P-PSDB; AAY43767, AAY43768, AAY43769, AAY43770, AAY43771, AAY43772,
FT DR AAY43773, AAY43774, AAY43775, AAY43776.
FT XX
FT PT New recombinant enzymes for biosynthesis of exopolysaccharides having
FT PT e.g. antitumor or probiotic properties or useful in fermented milk
FT PT products
FT XX
FT PS Claim 6; Page 77-97; 163pp: French.
FT XX
FT CC The present sequence represents the eps operon of Streptococcus
FT CC thermophilus Sfi139. The operon contains 10 open reading frames, and
FT CC encodes enzymes (eps1, eps2, eps3, eps4, eps5, eps6, eps7, eps8, eps9
FT CC and eps10) that are involved in the biosynthesis of exopolysaccharides
FT CC (EPS). The enzymes catalyse the formation of specific intersugar bonds.
FT CC The enzymes catalyse a process that includes at least one step of
FT CC forming a bond (in alpha or beta anomeric form) between C1, carrying
FT CC the reducing aldehyde group of an activated D-Galp (galactose in
FT CC pyranose form), and a phosphate group on a lipophilic or proteinaceous
FT CC primer. The enzymes are used to produce EPS that have antitumor or
FT CC probiotic properties or are used in foodstuffs to improve organoleptic
FT CC qualities and flavour. When expressed by lactic acid bacteria, EPS impart
```

```
CC a free-flowing character and/or a smooth, creamy texture to acidified
CC milk products (yoghurt or cheese).
XX
SQ Sequence 19966 BP; 6405 A; 3133 C; 3732 G; 6695 T; 1 other;
Query Match 2.7%; Score 80; DB 21; Length 19966;
Best Local Similarity 52.0%; Pred. No. 4.1e-06;
Matches 179; Conservative 0; Mismatches 165; Indels 0; Gaps 0;
QY 1307 ACATCTATAGAAGCTTTTACCAATAGAAGATTCACATATTCATAGAATACCTTTAGTTT 1366
FT ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 12794 ACTGTTAGAATATGATTTTATATAATTAGGAGTAGAATAAAGAGATCAATCCATTAAT 12853
QY 1367 CTATTTATATCCCGCTTATAACTGTGCAAAATTATATTCAAAGCATGTGTAGATGCTC 1426
FT ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 12854 CAATTATGTTCCACATATATATGTTGAAAATATATATTAGGACATGATTTGAATCAATCT 12913
QY 1427 TTAATCAAACTGTTGTCGATCTCGAGGTTTGTATTTTGAACGATGGTTCAACAGATAATA 1486
FT ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 12914 TAGCTCAAAACATATCGCAATATTGAAGTCATTATAGTAAATGATGTGACACAGATCAGT 12973
QY 1487 CCTTGAAGTGTATCAATAGCTTTATGTTGTAATAATCCTAGGGTACGCATCATCTCTAAAC 1546
FT ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 12974 CGCTAGCAGTAATTTCCGATTTAATCTGTAGTCATCATATAATTAAGGTAATCAACCAA 13033
QY 1547 CAAATGGCGGAATAGCCTCAGCATCAAAATGCAGCCGTTTCTTTTGTAAAGGTTATTACA 1606
FT ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13034 AAACCAAGGATATTCAGTAGCTCGAAACACTGGTATTGATGCGGCAACTGGTAATATA 13093
QY 1607 TTGGCAGTTAGATTCAGATGATTATCTTTGAGCCTGATGCAGTT 1650
FT ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13094 TAGCTTTTGTAGATGCAGATGACAAAATTAAGCCAGACTTTTGT 13137
RESULT 12
AAY43720/c
ID AAY43720 standard; DNA; 11474 BP.
XX
AC AAY43720;
XX
DT 22-DEC-2000 (first entry)
XX
DE LPS core biosynthesis locus.
XX
XX Biosynthetic locus; biosynthesis; lipid A biosynthesis;
KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAC transferase;
KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;
KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;
KW immunity; immunogen; ganglioside; ds.
XX
OS Campylobacter jejuni OH4384.
XX
FH Key Location/Qualifiers
FT CDS 350..1237
FT FT /*tag= a
FT FT /product= acyltransferase
FT FT /note= "Open reading frame 2b"
FT CDS 1234..2490
FT FT /*tag= b
FT FT /product= glycosyl transferase
FT FT /note= "Open reading frame 3a"
FT CDS 2786..3955
FT FT /*tag= c
FT FT /product= Glycosyl transferase
FT FT /note= "Open reading frame 4b"
FT CDS 4025..5068
FT FT /*tag= d
FT FT /product= Beta 1,4-GalNAC transferase
FT FT /note= "Open reading frame 5a"
FT CDS 5054..5959
FT FT /*tag= e
FT FT /product= Beta 1,2- galactosyltransferase
FT FT /note= "Open reading frame 6a"
```


QY 2532 CGCGATGATTTCTCAGCATTAACACATGATGGATCGAGAAAATCAATGCGCATCCACC 2591
DB 16467 TAGCACCACAACTAACCAATAAAGCAATCCGCTCTCTACTATAAACACACACAAAAA 16408
QY 2592 ATTTAAAGCTGATTAAACCTATTTTAATGACATGACTTAAAGAGTATGAATGTGAA 2651
DB 16407 TTAACGAACATATAACACACACCTATATCCCACTACTCAAAAACATAAATAAAT 16348
QY 2652 AGGGCATCACAGGTATGTTTATGAAGTATGCGGTACCGCATGAGCTTCTGACGATTAT 2711
DB 16347 TACTTAAACCCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 16288
QY 2712 TAAGAGTATCATCATCTGCGCAATCAATTGATAGTGGCAGATATAACACTCAGGA 2771
DB 16287 TAACCAACAAAACGAACTCTCAAAAATAAATAAATAAATAAATAAATAAATAA 16228
QY 2772 TATTTGGTTCCTAATTTGCACTTTTATCTTAGAAAAGAAACCGGCGATGTTTATAA 2831
DB 16227 ATATATCTCTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 16168
QY 2832 ACATCGACCTGACTTATATGCTTGGGACGAAATTAACATGACAAATGAACAAT 2891
DB 16167 AACCTTTACCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 16108
QY 2892 TCAAGTGCAAAAGGCGAAATATCCCGTTTAAACAAGTTTCATTTATTAATAGTATA 2951
DB 16107 TATCCCTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 16048
QY 2952 GCTATAAACAATTTGATTTTATAA 2979
DB 16047 CAAAAATATATCCCAATTTTATAACA 16020

RESULT 15
AAS63332/c
ID AAS63332 standard; DNA; 17527 BP.
XX
AC AAS63332;
XX
DT 29-JAN-2002 (first entry)

Chemically pretreated metabolism associated gene #27.

Human; cytostatic; anti-tumour; metabolism; metabolic disease; liver;
solid tumour; cancer; cytosine methylation; epigenetic; eye; kidney;
single nucleotide polymorphism detection; SNP; stool; urine; lung;
cerebral-spinal fluid; intestine; brain; heart; prostate; breast;
DUSP2; EPHX2; QDPR; SGSH; SHMT2; SLC7A2; SLC7A4; TYMS; ds.

Homo sapiens.

WO2001/76451-A2.

18-OCT-2001.

06-APR-2001; 2001WO-EP04016.

06-APR-2000; 2000DE-1019058.

07-APR-2000; 2000DE-1019173.

30-JUN-2000; 2000DE-1032529.

01-SEP-2000; 2000DE-1043826.

(EPIG-) EPIGENOMICS AG.

Olek A, Piepenbrock C, Berlin K;

WPI; 2002-010834/01.

New nucleic acid, useful for diagnosis and therapy of metabolic
disease, solid tumour and cancers, comprises segment of chemically
modified genomic sequences of genes associated with metabolism

Claim 1; Page 79-84; 143pp; English.

XX

CC The invention relates to a nucleic acid (I) comprising a sequence at
CC least 18 bases of a segment of the chemically pretreated DNA of genes
CC associated with metabolism such as DUSP2 (NM_004418), EPHX2 (NM_001979),
CC QDPR (NM_000320), SGSH (NM_000199), SHMT2 (NM_005412), SLC7A2
CC (NM_003046), SLC7A4 (NM_004173) and TYMS (NM_001071) (all
CC undefined). (I) are useful for diagnosis and therapy of metabolic
CC disease, solid tumours and cancers; as primer oligonucleotides for the
CC amplification of DNA sequences, for detecting the cytosine methylation
CC state and/or single nucleotide polymorphisms (SNPs) in a chemically
CC treated DNA of genes associated with metabolism. An array of (I) is
CC useful for ascertaining genetic and/or epigenetic parameters for the
CC diagnosis and/or therapy of existing diseases or the predisposition to
CC specific diseases by analysing cytosine methylations. The method involves
CC chemically treating genomic DNA sample by a solution of bisulphite,
CC hydrogen sulphite or disulphite such that cytosine bases which are
CC un methylated at the 5th position are converted to uracil or another base
CC which is dissimilar to cytosine in terms of hybridisation behaviour and
CC amplifying fragments of the chemically pretreated genomic DNA. The
CC genomic DNA is from cells or cellular components which contain DNA,
CC sources of DNA comprising, for e.g. cell lines, biopsies, blood, sputum,
CC stool, urine, cerebral-spinal fluid, tissue embedded in paraffin such as
CC tissue from eye, intestine, kidney, brain, heart, prostate, lung, breast
CC or liver, histologic object slides and their combinations. Genetic
CC parameters are mutations, in particular insertions, deletions, point
CC mutations, inversions and polymorphisms of genes associated with
CC metabolism and sequences further required for their regulation.
CC Epigenetic parameters are in particular cytosine methylations and
CC further chemical modifications of DNA bases of genes associated with
CC metabolism. Further epigenetic parameters include for e.g. the
CC acetylation of histones which correlates with DNA methylation.
CC AAS63306-AAS63373 represent chemically pretreated metabolism associated
CC genes, and related primers of the invention.

XX SQ Sequence 17527 BP; 4501 A; 296 C; 4094 G; 8636 T; 0 other;

Query Match 2.68; Score 77.2; DB 24; Length 17527;
Best Local Similarity 43.9%; Pred. No. 1.4e-05;
Matches 434; Conservative 0; Mismatches 543; Indels 11; Gaps 2;

QY 1992 AAGAGAAAAACCATTTTGTGTAGTCAATCAGTCATTAATAAGACAAAGGCATCAATTA 2051
DB 16996 ATAAAAAAGGTATCTTAACTATAAACATTTAAAAAATACTAAACCAACAGCA 16937
QY 2052 TAATATTGACAAATTTGATGATTAGATGAAAGTAGAAGTATATCTTCAATAAACCCG 2111
DB 16936 TAATACACCTTAAATCCCACTACTCAAAAACAAAAATAAATAATTTACTTAAACCC 16877
QY 2112 TGAATATCAAGAAGAAATGGATATGTTAAAAGATCTTAAACTCATTTAAAAATAAGATGC 2171
DB 16876 AAAATTTAAATAACATAAACCAAAATCACACCCCTATCTCCAACTTAACACACAAA 16817
QY 2172 CAAAATCGCAGTCAGTATTTTCTATCCCAATACATTAACCGCTTAGTGAAAAAACTAAA 2231
DB 16816 CAAAACCCCACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 16757
QY 2232 CAATATTATGAATATAATAAATAATATTCGTTTATTCTTACATGTTGATAAGATCA 2291
DB 16756 AAAAAACAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 16701
QY 2292 TCTTACACGACATCAAAAAAGAAATATTGGCTTTCTATCATAGACCAAGTGAATAT 2351
DB 16700 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 16641
QY 2352 TTTACTAAATAATGACATCTCATATTACACGAGTAAATAGACTTAATAAATCGAGGCACA 2411
DB 16640 TTTCTAAATAATACGACTCAATAAATAAATAAATAAATAAATAAATAAATAAATAA 16588
QY 2412 TTTAAGTAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2471
DB 16587 TTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 16528
QY 2472 TAATCATGACACGCTTTCGTTTAAAAAATGACAGCTATGCTTTATGAAAAAATAATGATGT 2531

```

Db 16527 TACCTATAATCCTAACACTTTTAAAAAATCAAAACAAATAAATCACCTAAAAATCAAAAAAT 16468
Qy 2532 CGGCATGAATTTCTCAGCATTAAACACATGATTGGATCCAGAAAAATCAATGGCATCCACC 2591
Db 16467 TACGACCAACTAACCAAAATACGAAATCCGGTCTCTACTAAAAACACACACAAAAAA 16408
Qy 2592 ATTTAAAAAGCTGATTAAACCTATTTTAAATGACAATGACTTAAGAAGTATGAATGTGAA 2651
Db 16407 TTAACGAAACATAATAACACACACCTATAATCCCAACTACTCARAAAACTAAAAATAAAT 16348
Qy 2652 AGGGCATCACAAGTATGTTTATGAAGTATGCGCTACCGCATGAGCTTCTGACGATTAT 2711
Db 16347 TACTTAAACCCCAAAAAATAAAAAATTTACTATAAACCTAAATCCTACCATTAACACTCAACC 16288
Qy 2712 TAAAGAAGTCATCACATCCTGCCAATCAATTGATAGTGTGCCAGAAATATAACACTGAGGA 2771
Db 16287 TAAACACAAAACGAAACTCCTCAAAAAAATTAATTAATTTAATTTAAAAA 16228
Qy 2772 TATTTGGTTCCAATTTGCACCTTTTAAATCTTAGAAAAAGAAACCGGCCATGTATTTAATAA 2831
Db 16227 ATATAATCTCTATATAAAATATAAAATAAAAATACAAAAAATATAATTTATAAAATAA 16168
Qy 2832 AACATCGACCTGACTTATATGCCTTGGGAACGAAATTAACAATGGACAAAATGAACAAAT 2891
Db 16167 AACCTTTACCAATAAAAAACCAAAAAACAAAAAACTATAAACGAAAAATTTACTACTA 16108
Qy 2892 TCAAGTCAAAAAAGCGGAAATATCCCGTTAACAAAGTTCAATTATAGTATTAAC 2951
Db 16107 TATCCCTATTAAAAATCAAAACTATCTCTCATATATTCCTTCAAAACATAACTATAAATTA 16048
Qy 2952 GCTATAAAACATTTGCATTTTATTAAAA 2979
Db 16047 CAAAAATATTATCCCAATTTTATAACA 16020

```

Search completed: January 3, 2003, 21:32:20
 Job time : 737 secs

THIS PAGE BLANK (USPTO)

BEST AVAILABLE COPY

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2003, 21:18:18 ; Search time 117 Seconds
(without alignments)
7808.462 Million cell updates/sec

Title: US-09-842-484a-1

Perfect score: 2979

Sequence: 1 ttataactgattaaagaag.....acatttcattttattaaaa 2979

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA.*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*

5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2979	100.0	2979	4	US-09-437-277-4
2	1503.2	50.5	2112	4	US-09-437-277-2
3	74.6	2.5	14602	1	US-08-597-236-1
4	74.6	2.5	14602	1	US-08-746-682A-1
5	74.2	2.5	615	4	US-08-998-416-186
6	70.6	2.4	837	4	US-08-998-416-288
7	70.2	2.4	636	4	US-08-998-416-1137
8	67.2	2.3	32768	4	US-08-961-527-71
9	66.8	2.2	1956	4	US-08-559-896B-1
10	63.2	2.1	19124	2	US-08-487-826B-13
11	61	2.0	19390	4	US-08-961-527-86
12	60.6	2.0	2023	3	US-08-961-083-199
13	58.8	2.0	658	4	US-08-998-416-595
14	57.8	1.9	6243	2	US-08-056-075-1
15	57.6	1.9	811	3	US-08-961-083-197
16	55.6	1.9	7218	1	US-08-232-463-14
17	55	1.8	14187	4	US-09-453-702B-121
18	54.2	1.8	5852	1	US-07-867-106-2
19	53.8	1.8	663	4	US-08-998-416-191
20	53.6	1.8	8920	2	US-08-446-855A-1
21	53.6	1.8	8920	4	US-09-150-741-1
22	52.2	1.8	854	4	US-08-998-416-534
23	52.2	1.8	168575	4	US-09-426-290-1
24	51.6	1.7	724	4	US-08-998-416-683
25	51	1.7	711	4	US-08-998-416-786
26	51	1.7	3095	6	5231168-1
27	51	1.7	3480	1	US-08-920-812-11

28	51	1.7	3480	1	US-08-920-827-11	Sequence 11, Appl
29	51	1.7	3480	1	US-08-921-177-11	Sequence 11, Appl
30	51	1.7	3480	1	US-08-362-577C-11	Sequence 11, Appl
31	51	1.7	3480	2	US-08-920-828-11	Sequence 11, Appl
c 32	50.4	1.7	1422	2	US-08-319-704-5	Sequence 5, Appl1
c 33	50.2	1.7	19124	2	US-08-487-826B-13	Sequence 13, Appl1
34	49.8	1.7	1431	4	US-09-316-083-2	Sequence 2, Appl1
35	48.8	1.6	827	4	US-08-998-416-535	Sequence 535, App
36	48.8	1.6	5181	1	US-08-257-073-10	Sequence 10, Appl
37	48.6	1.6	660	1	US-07-991-867B-32	Sequence 32, Appl
38	48.6	1.6	660	1	US-08-107-755A-32	Sequence 32, Appl
39	48.6	1.6	660	2	US-08-544-332-32	Sequence 32, Appl
40	48.6	1.6	660	4	US-09-370-861A-32	Sequence 32, Appl
41	48.6	1.6	1511	1	US-07-991-867B-8	Sequence 8, Appl1
42	48.6	1.6	1511	1	US-08-107-755A-8	Sequence 8, Appl1
43	48.6	1.6	1511	2	US-08-544-332-8	Sequence 8, Appl1
44	48.6	1.6	1511	4	US-09-370-861A-8	Sequence 8, Appl1
45	48.6	1.6	4507	2	US-08-568-459A-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1

US-09-437-277-4

; Sequence 4, Application US/09437277

; Patent No. 6444447

; GENERAL INFORMATION:

; APPLICANT: THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA

; TITLE OF INVENTION: POLYMER GRAFTING BY POLYSACCHARIDE SYNTHASES

; FILE REFERENCE: 5820.551

; CURRENT APPLICATION NUMBER: US/09/437,277

; CURRENT FILING DATE: 1999-11-10

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: WordPerfect 8.0 (saved in ASCII format)

; SEQ ID NO 4

; LENGTH: 2979

; TYPE: DNA

; ORGANISM: Pasteurella multocida

US-09-437-277-4

Query Match 100.0%; Score 2979; DB 4; Length 2979;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2979; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TTATAA	CTGATTAAAGAGGTAACGATTCACGATTCACGAGGTTAAATTTTAAAGGAAAGAAA	60
Db	1	TTATAA	CTGATTAAAGAGGTAACGATTCACGATTCACGAGGTTAAATTTTAAAGGAAAGAAA	60
Qy	61	ATGAAT	ACATTTATCACAAGCAATAAAGCATATACAGCAATGACTATGATTTAGCACTC	120
Db	61	ATGAAT	ACATTTATCACAAGCAATAAAGCATATACAGCAATGACTATGATTTAGCACTC	120
Qy	121	AAATTAT	TTTGGAGAAGTCTGCTGAAACCTACGGCGGAAAAATTCGTTGAAATTCCTCAAAATATTC	180
Db	121	AAATTAT	TTTGGAGAAGTCTGCTGAAACCTACGGCGGAAAAATTCGTTGAAATTCCTCAAAATATTC	180
Qy	181	AAATGTA	AGAAAGAACTCTCGACCAATTTCTTATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGT	240
Db	181	AAATGTA	AGAAAGAAAGCTCTCGACCAATTTCTTATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGT	240
Qy	241	TGCGAT	AGCTCATTTAGATATCGCAACACAGCTCTTACTTTCCAACTGTAATAAATAAATAAATTA	300
Db	241	TGCGAT	AGCTCATTTAGATATCGCAACACAGCTCTTACTTTCCAACTGTAATAAATAAATAAATTA	300
Qy	301	CTATCC	GAATCAGAAAAAACAGTTTAAAAATAAATGGAATCTATCTACTTGGGAAAAAAA	360
Db	301	CTATCC	GAATCAGAAAAAACAGTTTAAAAATAAATGGAATCTATCTACTTGGGAAAAAAA	360
Qy	361	TCGGAG	AACCGAGAAATCAGAAAGTGGAACTAGTACCCAAAGATTTTCTCTAAAGATCTTT	420
Db	361	TCGGAG	AACCGAGAAATCAGAAAGTGGAACTAGTACCCAAAGATTTTCTCTAAAGATCTTT	420

Qy	421	GTCTCTGCTCCATTCGCCAGATCATGTTAAATGATTTTACATGGTACAAAATTCGAAAAA	480
Db	421	GTCTCTGCTCCATTCGCCAGATCATGTTAAATGATTTTACATGGTACAAAATTCGAAAAA	480
Qy	481	AGCTTAGGTATAAGCCGTGAATAAGAAATATCGGTCTTCTCTATATTTATTTCTACATTT	540
Db	481	AGCTTAGGTATAAGCCGTGAATAAGAAATATCGGTCTTCTCTATATTTATTTCTACATTT	540
Qy	541	AATCGTACCCGTATTTTAGATATACGTTTACGCTGTTTGGTCAATTCAGAAAACAAACTAC	600
Db	541	AATCGTACCCGTATTTTAGATATACGTTTACGCTGTTTGGTCAATTCAGAAAACAAACTAC	600
Qy	601	CAATTTGAAGTCGTTTGGCAGATGATGGTATAGGAAAACTTACTTACCATTTGTGCAA	660
Db	601	CAATTTGAAGTCGTTTGGCAGATGATGGTATAGGAAAACTTACTTACCATTTGTGCAA	660
Qy	661	AAATACGAACAAAACCTTGACATAAAGTATGTAAGACAAAAAGATATGGGATATCAATTTG	720
Db	661	AAATACGAACAAAACCTTGACATAAAGTATGTAAGACAAAAGATATGGGATATCAATTTG	720
Qy	721	TGTGCAGTCAGAAACTTAGGTTTACGTACAGCAAAAGTAGATTTTGTCTCGATCTAGAC	780
Db	721	TGTGCAGTCAGAAACTTAGGTTTACGTACAGCAAAAGTAGATTTTGTCTCGATCTAGAC	780
Qy	781	TCCGATATGGCACCAACAATATATGCGTTCATCTTCTTACAGAACTATTAGAAGC	840
Db	781	TCCGATATGGCACCAACAATATATGCGTTCATCTTCTTACAGAACTATTAGAAGC	840
Qy	841	AATGATATTTGTTTAAATTTGGACCTAGAAAATATGTGGACTCATATATATCCGCAGAA	900
Db	841	AATGATATTTGTTTAAATTTGGACCTAGAAAATATGTGGACTCATATATATCCGCAGAA	900
Qy	901	CAATTCCTTAACGATCCATATTTAATAGAACTACTCTGAAACCGCTACAATACAAAT	960
Db	901	CAATTCCTTAACGATCCATATTTAATAGAACTACTCTGAAACCGCTACAATACAAAT	960
Qy	961	CTTCTCGATTACATCAAAAGGAATATATCGTTGGATGGAGATTAGAACATTTCAAAAA	1020
Db	961	CTTCTCGATTACATCAAAAGGAATATATCGTTGGATGGAGATTAGAACATTTCAAAAA	1020
Qy	1021	ACGGATAACTAGCTCTATGTAATCTCGTTTCGTTATTTTGTTCGGGTATGTGCA	1080
Db	1021	ACGGATAACTAGCTCTATGTAATCTCGTTTCGTTATTTTGTTCGGGTATGTGCA	1080
Qy	1081	TTTTCTAAAAGATGGCTAAATAAGTAGGTGGTTCGATCAAGAAATTAATCATTTGGGG	1140
Db	1081	TTTTCTAAAAGATGGCTAAATAAGTAGGTGGTTCGATCAAGAAATTAATCATTTGGGG	1140
Qy	1141	GGCGAAGATGATAATTTGGTTACAGATTAATTTGCCAAAGCGTGTTTTTCAGAGTAAT	1200
Db	1141	GGCGAAGATGATAATTTGGTTACAGATTAATTTGCCAAAGCGTGTTTTTCAGAGTAAT	1200
Qy	1201	GACGGCGGAATGCCATCCATCAAGAACACCTGGTAAAGAAAATGAACAGAACCGCAA	1260
Db	1201	GACGGCGGAATGCCATCCATCAAGAACACCTGGTAAAGAAAATGAACAGAACCGCAA	1260
Qy	1261	GCTGGTAAAAAGTATTACGCTTAAATTTGTGAAGAAAAGGTACTCTACATCTATAGAAG	1320
Db	1261	GCTGGTAAAAAGTATTACGCTTAAATTTGTGAAGAAAAGGTACTCTACATCTATAGAAG	1320
Qy	1321	CTTTTACCAATAGAAGATCCACATATTCATAGAAATACCTTTAGTTTCTATTTATATCCCC	1380
Db	1321	CTTTTACCAATAGAAGATCCACATATTCATAGAAATACCTTTAGTTTCTATTTATATCCCC	1380
Qy	1381	GCTTTAACTGTGCAAAATTAATTTCAAGATGTTAGATAGTGCTCTTAATCAAACTGTT	1440
Db	1381	GCTTTAACTGTGCAAAATTAATTTCAAGATGTTAGATAGTGCTCTTAATCAAACTGTT	1440
Qy	1441	GTCGATCTCAGGTTGTTATTTGTACAGTAGGTCTACAGATAATACCTTAGAAGTGATC	1500
Db	1441	GTCGATCTCAGGTTGTTATTTGTACAGTAGGTCTACAGATAATACCTTAGAAGTGATC	1500
Qy	1501	AATTAAGCTTTATGGTAATAATCTTAGGGTAGCGATCAATGCTTAACCAAAATGCGGAA	1560

[illegible]

Db 2581 GCACATCCACCATTAAAAAGCTGATTAAAAACCTATTATTAATGACAATGACTTAAGAAGT 2640
Qy 2641 ATGAATGTGAAGGGGCATCACAGGTATGTTTATGAAGTATGCGCTACCGCATGAGCTT 2700
Db 2641 ATGAATGTGAAGGGGCATCACAGGTATGTTTATGAAGTATGCGCTACCGCATGAGCTT 2700
Qy 2701 CTGACGATTTAAAGAAGTCATCACATCTGCGCAATCAATTTGATAGTGTGCCAGATAT 2760
Db 2701 CTGACGATTTAAAGAAGTCATCACATCTGCGCAATCAATTTGATAGTGTGCCAGATAT 2760
Qy 2761 AACACTGAGGATTTTGGTTCCTAATTTGACATTTTATATCTAGAAAAGAAACCGGCAT 2820
Db 2761 AACACTGAGGATTTTGGTTCCTAATTTGACATTTTATATCTAGAAAAGAAACCGGCAT 2820
Qy 2821 GTATTTAAATAACATCGACCTGACATATATGCTTTGGGAACGAAATTAACAATGACA 2880
Db 2821 GTATTTAAATAACATCGACCTGACATATATGCTTTGGGAACGAAATTAACAATGACA 2880
Qy 2881 AATGAACAAATTCAAAGTGCAGAAAGGGCGAAATATCCCGGTTAACAGTTTCATTATT 2940
Db 2881 AATGAACAAATTCAAAGTGCAGAAAGGGCGAAATATCCCGGTTAACAGTTTCATTATT 2940
Qy 2941 AATAGTATACGCTATATAAATTTGATTTTATATAAA 2979
Db 2941 AATAGTATACGCTATATAAATTTGATTTTATATAAA 2979

RESULT 2
US-09-437-277-2
; Sequence 2, Application US/09437277
; Patent No. 644447
; GENERAL INFORMATION:
; APPLICANT: THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA
; TITLE OF INVENTION: POLYMER GRAFTING BY POLYSACCHARIDE SYNTHASES
; FILE REFERENCE: 5820.551
; CURRENT APPLICATION NUMBER: US/09/437, 277
; CURRENT FILING DATE: 1999-11-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: WordPerfect 8.0 (saved in ASCII format)
; SEQ ID NO 2
; LENGTH: 2112
; TYPE: DNA
; ORGANISM: Pasteurella multocida
US-09-437-277-2

Query Match 50.5%; Score 1503.2; DB 4; Length 2112;
Best Local Similarity 82.5%; Pred. No. 0;
Matches 1743; Conservative 0; Mismatches 348; Indels 21; Gaps 1;

Qy 61 ATGAATACATTTATCACAGCAATAAAAAGCATATAACAGCAATGACTATGAATTAGCACTC 120
Db 1 ATGAATACATTTATCACAGCAATAAAAAGCATATAACAGCAATGACTATCAATTAGCACTC 60
Qy 121 AAATTTATTTGAGAAGTCTGCTGAACCTACGGCGGAAATAATCGTTGAATCCAAATATAC 180
Db 61 AAATTTATTTGAAAAGTGGCGGAAATCTATGGACGGAATAATGTTGAATTTCAAATACC 120
Qy 181 AAATGTAAGAAAACCTCTCGACCAATTC-----TTATGTRAGT 219
Db 121 AAATGCCAAGAAAACCTCTCAGCACATCTCTGTTAATTCAGCACATCTTTTCTGTAAAT 180
Qy 220 GAAGATAAAAAAAGTGTGCGATAGCTCAATAGATATCGCAACACAGCTCTTACTT 279
Db 181 AAAGAAGAAAAGTCAATGTTGCGATAGTCCGTTAGATATGCAACACACTGTTACTT 240
Qy 280 TCCAACGTAAAAAATTAACCTCTATCCGAATCAGAAAAAAGACAGTTTAAAAAATAAATGG 339
Db 241 TCCAACGTAAAAAATTAAGTACTTCTGACTCGGAAAAAAGACAGTTTAAAAAATAAATGG 300
Qy 340 AAATCTATCTACTGGGAAAAAATTCGGAGAACCGCAGAAATTCAGAAAGGTGGAATAGTACC 399
Db 301 AAATGCTCTACTGAGAAAGAAATCTGAAATGCGGAGGTGAAGCGGCTGCCCTTGTACCA 360

Qy 400 AAAGATTTTCCCTAAAGATCTTGTCTTCCATTCATCCAGATCATGTTAAATGATTTTACA 459
Db 361 AAAGATTTTCCCAAGATCTGCTTTTAGCGCTTTACCTGATCATGTTAAATGATTTTACA 420
Qy 460 TGGTACAAAAATCGAAAAAAGCTTAGGTATAAAGCCCTGTAAATTAAGAAATATCGGTCCT 519
Db 421 TGGTACAAAAAAGCGAAAGAAAGACTTGGCATATAAACCTGAACATCAACATCTTGGTCCT 480
Qy 520 TCTATTATTATTCCCTACATTTAAATCGTAGCCGATATTTTAGATATAACGTTTACCCCTGTTTG 579
Db 481 TCTATTATCTGTACAAACATTTCAATCGACCAACAAATTTTATCGATTTACATACCCCTGTTTA 540
Qy 580 GTCATATCAAAAAAACAATACCCATTTGAAGTCGTTGTTGACAGATGATGGTAGTAAGGAA 639
Db 541 GTAAACCAAAAAACACATTTACCCGTTTGAAGTTATCGTGACAGATGATGGTAGTCAGGAA 600
Qy 640 AACTTACTTACCATTTGTCAAAAATACGAACAAAACCTTGACATAAAGTATGTAAGACAA 599
Db 601 GATCTATCACCGATCATTTCCGCAATATGAAAATAAATTTGGATATTTCGCTACGTCAGACAA 660
Qy 700 AAAGATTTATGGATATCAATTTGTGTCAGTTCAGAAACCTTAGGTTTACGTACACGCAAGTAT 759
Db 661 AAAGATAACGGTTTCAAGCCAGTCCCGCTCGGAATATGGGATTTACGCTTACGCAAAATAT 720
Qy 760 GATTTTGTCTCGATTTCTAGACTGCGATATGGCACCACACAAATTTATGGGTTTCATTTCTTAT 819
Db 721 GACTTTATTGGCTTACTCGACTGTGATATGGCGCAAAATCCCATTTATGGGTTTCATTTCTTAT 780
Qy 820 CTTACAGAACTATTAGAAGACAAATGATATTGTTTAAATTTGGACCTAGAAAAATATGTGGAT 879
Db 781 GTTCGAGAGCTATTAGAAGATGATGATTTAAACAATCATTTGGTCCAAAGAAAATACATCGAT 840
Qy 880 ACTCATATATTACCGCAGAACAAATTCCTTAAAGATCCATATTTAAATGAACATCACACT 939
Db 841 ACACAAATATTGACCCCAAGAGACTTCTTAAATAAGCGAGTTGCTTGAATCAATTAACCA 900
Qy 940 GAAACCGCTACAAAATAACATCTCTCGATTTACATCAAAAGGAAATATATCGTTGGATGG 999
Db 901 GAAGTGAAACCAATAATAGTGTCCCGCAAAAGGGGAGGAACAGTTTCTCTGGATTTGG 960
Qy 1000 AGATTTAGAACATTTCAAAAAACCGATATCTACGCTATGATGATTCCTCGGTTTCGTTAT 1059
Db 961 CGCTTTAGAACAAATTCGAAAAACAGAAAATCTCCGCTTATCCGATTCGCGTTCCTT 1020
Qy 1060 TTTGTTGCGGGTAAATGTTGCATTTTCTAAAGAAATGGCTAAATAAGTAGTGGTTCGAT 1119
Db 1021 TTTGCGCGGGTAAATGTTGCTTTCCGCTAAAAAATGGCTAAATAAGTAGTGGTTCCTTAT 1080
Qy 1120 GAAGAAATTTAATCAATTTGGGGGGGGAAGATGTAGAAATTTGGTTACAGATTTATTTGCCAAA 1179
Db 1081 GAGGAATTTAATCACTCTGCGGTGAGAGATGTGGAATTTGGATATCGCTTATTTCCGTTAC 1140
Qy 1180 GCCTGTTTTTCAGAGTAAATTCACGCGGGAATGGCCATCCATCAAGAACCCCTGGTAAA 1239
Db 1141 GGTAGTTTTCTTAAACATTTGATGGCATTTATGGCTATTTGGCTCGATATTATGAGAGAAAAG 1200
Qy 1240 GAAATTAAGACAGACGCAAGCTGTGTAAGAACTATTACGCTTAAATTTCTGAAAGAAAAG 1299
Db 1201 GAAATTAAGACCGATCGTGAGCGGGAATAATATTACGCTCGATATTATGAGAGAAAAG 1260
Qy 1300 GTACCTTACATCTATAGAAAAGCTTTTACCAATAGAAGATTACATATTTCATAGAATACCT 1359
Db 1261 GTCCCTTATCTATAGAAAACCTTTTACCAATAGAAGATTTCGCATATCAATAGAGTACCT 1320
Qy 1360 TTAGTTTCTATTATTATATCCCGCTTATAACTGTGCAAAATATTATTTCAAAAGATGTGTAGAT 1419
Db 1321 TTAGTTTCAAAATTTATATCCAGCTTATAACTGTGCAAACTATATTTCACAGTTCGCTAGAT 1380
Qy 1420 AGTGTCTTAAATCAACACTGTTGCTCGATCTCGAGGTTTGTATTGTAACCATGTTTCAACA 1479
Db 1381 AGTGCACGAAATCAGACTGTTGTTGCTCGAGGTTTGTATTGTAACCATGTTTCAACA 1440
Qy 1480 GATAATACCTTAGAAGTATCAATAAGCTTTTATGGTAAATAATCCTTAGGTCAGGCATCATG 1539

FEATURE:
NAME/KEY: terminator
LOCATION: 230..252
FEATURE:
NAME/KEY: promoter
LOCATION: 274..302
FEATURE:
NAME/KEY: RBS
LOCATION: 340..345
US-08-597-236-1

Query Match 2.5%; Score 74.6; DB 1; Length 14602;
Best Local Similarity 50.7%; Pred. No. 1e-07;
Matches 179; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

QY 1343 ATATTCATAGAACCTTTAGTTCTTATTTATATCCCGCTTATACTGTCRAATTATA 1402
DB 8219 AATGTATCTTAAAGTCTATCTCTATTGTTATTCAGTATATAATGTAGAGAAATATT 8278
QY 1403 TTCAAGATGTAGATAGTCTCTTAATCAAACTGTTGTCGATCTCGAGGTTTGTATTT 1462
DB 8279 TAGAAAAATCTTTCGAATCTGTTCAAAATCAGACTTACAATAATTTGAAGTGATTAG 8338
QY 1463 GTAAGATGTTCAACAGATAAATACCTTAGAAGTGATCAATAGCTTTATGGTAATATC 1522
DB 8339 TGAATGATGCTCAACCGGATTCATCACTTTCAATATCGGAAAAATTTGTTAATCAGGATA 8398
QY 1523 CTAGGTACGCATCATCTCTAAACCAATGCGGAATAGCCTCAGCATCAATTCAGCCG 1582
DB 8399 AAGATTTTCTGTTTCTTAAAGAAATGGTGATGTGTCATCTGCACGAAATTTTGGAA 8458
QY 1583 TTTCTTTTGTCTAAGGTTTATACATTTGGCAGTTAGATTTCAGATGATATCTTGAGCCTG 1642
DB 8459 TTAAGAGGCTAAGGATCGTTTATCACTTTGTAGATAGTATGATGATACATAGTAAAG 8518
QY 1643 ATGAGTTGAACGTGTTTAAAAAGAAATTTTAAAGATAAAACGCTAGCTTGT 1695
DB 8519 ATTATCTTCTCATTTGGTAGTGGGATAAAAGTGAGACCTCTATAGTTTGT 8571

RESULT 4

US-08-746-682A-1
Sequence 1, Application US/08746682A
Patent No. 5786184
GENERAL INFORMATION:
APPLICANT: STINGELE, Francesca
APPLICANT: MOLLET, Beat
TITLE OF INVENTION: LACTIC BACTERIA PRODUCING
TITLE OF INVENTION: EXOPOLYSACCHARIDES
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americans
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/746,682A
FILING DATE: 14-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/597,236
FILING DATE: 20-JUN-1995
APPLICATION NUMBER: EP 95201669.9
FILING DATE: 20-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fanucci A., Allan

REGISTRATION NUMBER: 30256
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 14602 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 352..1803
OTHER INFORMATION: /product= "epsA"
FEATURE:
NAME/KEY: CDS
LOCATION: 1807..2535
OTHER INFORMATION: /product= "epsB"
FEATURE:
NAME/KEY: CDS
LOCATION: 2547..3239
OTHER INFORMATION: /product= "epsC"
FEATURE:
NAME/KEY: CDS
LOCATION: 3249..3995
OTHER INFORMATION: /product= "epsD"
FEATURE:
NAME/KEY: CDS
LOCATION: 4051..4731
OTHER INFORMATION: /product= "epsE"
FEATURE:
NAME/KEY: CDS
LOCATION: 4898..5854
OTHER INFORMATION: /product= "epsF"
FEATURE:
NAME/KEY: CDS
LOCATION: 6425..7540
OTHER INFORMATION: /product= "epsG"
FEATURE:
NAME/KEY: CDS
LOCATION: 7736..8212
OTHER INFORMATION: /product= "epsH"
FEATURE:
NAME/KEY: CDS
LOCATION: 8221..9192
OTHER INFORMATION: /product= "epsI"
FEATURE:
NAME/KEY: CDS
LOCATION: 9285..10364
OTHER INFORMATION: /product= "epsJ"
FEATURE:
NAME/KEY: CDS
LOCATION: 10392..11339
OTHER INFORMATION: /product= "epsK"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 11302..12222
OTHER INFORMATION: /product= "CDS (eps L) covering CDS
OTHER INFORMATION: (eps k) on nucleotides 10392-11339"
FEATURE:
NAME/KEY: CDS
LOCATION: 12233..13651
OTHER INFORMATION: /product= "epsM"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 13732..14305
OTHER INFORMATION: /function= "CDS on the
OTHER INFORMATION: complementary strand"
FEATURE:
NAME/KEY: terminator
LOCATION: 230..252
NAME/KEY: promoter
LOCATION: 274..302

Mon Jan 6 18:39:49 2003

us-09-842-484a-1.rni

FEATURE:
NAME/KEY: RBS
LOCATION: 340..345
US-08-746-682A-1

Query Match 2.5%; Score 74.6; DB 1; Length 14602;
Best Local Similarity 50.7%; Pred. No. 1e-07;
Matches 179; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

QY 1343 ATATTCATAGAAATACCTTTAGTCTTATATATCCCGCTTATAACATGTCGAATATA 1402
DB 8219 AAATGTATCTTAAAGTCTAATCTCTATGTATATCCAGTATATAATGTAGAGAAATATT 8278
QY 1403 TTCAAAGATGTAGATAGTCTCTTAAATCAACCTGTTGTCGATCTCGAGGTGTGTAATTT 1462
DB 8279 TAGAAAAATGTTTGAATCTGTTCAAAATCAGACTTACAAATTTTGAAGTGAATTTAG 8338
QY 1463 GTAAGATGGTTTCAACAGATAATACCTTTAGAGTGATCAATAGCTTTATGGTAATATC 1522
DB 8339 TGAATGATGGCTCAACCGGATCATCTTCAATATGCGAAAAATTTGTTATCAGGATA 8398
QY 1523 CTAGGTAGCATCATGTCTAAACCAATGCGGAATAGCCCTCAGCATCAATGCGCGG 1582
DB 8399 AAAGATTTTCTGTTTCTTAAAGAAATGCGGTATGTCATCTGCACGAATTTTGGNA 8458
QY 1583 TTTCTTTTCTTAAAGTATTACATTTGGGCAGTTAGATTTCAGATGATTATCTTTGAGCCCTG 1642
DB 8459 TTAAGAAGCTAAGAGATCGTTTATCACATTTGTAGATGATGACTACATAGTAAAG 8518
QY 1643 ATGAGTGAATGCTTTTAAAGAAATTTTAAAGATATAAAGCGTAGCTTCT 1695
DB 8519 ATTATCTTTCTCATTTGGTAGTGGGATAAAAGTGAGACCTCTATAGTTTGT 8571

RESULT 5

US-08-998-416-186/c
; Sequence 186, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 186:
SEQUENCE CHARACTERISTICS:
LENGTH: 615 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE: PAG1074RP
ORGANISM: PAG1074RP
US-08-998-416-186

Query Match 2.5%; Score 74.2; DB 4; Length 615;
Best Local Similarity 46.6%; Pred. No. 4.8e-08;
Matches 238; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

QY 2019 TCAGTCATTAAATAGACAGGCGATCAATTTATTAATTAATGACAAATTTTGATGATTAGA 2078
DB 551 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 492
QY 2079 TGAAGTAGAAGATATATCTTCAATAAAACCGCTGAATATCAAGAGAAGAAATGGATGTT 2138
DB 491 ATAAGAAATTTAAAGTTAAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 432
QY 2139 AAAAGATCTTAAACTCAATTCAAAATAAAGATGCCAAATCGCAGTCAGTATTTCTATCC 2198
DB 431 AAATCAACATAATATTTATATAAATAGATATTTATATAAATAAATAATTAATTAATTAATTT 372
QY 2199 CAATACATTTAAAGCGCTTAGTGAAAAAAGCTTAAACATAATATTTTGAATATATAATAAATAT 2258
DB 371 AAATAAATTTAAATCTTTTATAATAATAAATAATATTTTAAATAAACAATTAATAATAA 312
QY 2259 ATTCGTTATTTCTACATGTTGATAAGATCATCTTACACAGACATCAAAAGAAAT 2318
DB 311 TATATTAATTAATGATAATCTATTTAATAATTTTAAAGAAATAATATATATATCTAATA 252
QY 2319 ATGGCTTTCTATCATAAGCAGCAAGTGAATATTTTACTAAATAAGTACATCATATTA 2378
DB 251 TATTTTAATAACTAATTTTAAATTTTGAACATAGACATAATAGTATTCATATAATAATTA 192
QY 2379 CAGAGTAATAGACTAATAAAAGCTGAGGACACATTTAAGTAATATTAATAAATAATTAAT 2438
DB 191 TTTTATAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 132
QY 2439 GTTAAATCTAAATTTGTAATACATCTTTTGTGATAATCATGACAGCCTATTCGTTAAAAA 2498
DB 131 AATAAGTATTAAATAAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 72
QY 2499 TGACAGCTATGCTTATATGAAAAAATATGAT 2529
DB 71 TTTAATTAATTAATTAATTAATAAAAAAGTT 41

RESULT 6

US-08-998-416-288/c
; Sequence 288, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park

STATE: No. 6239264th Carolina
COUNTRY: USA
Zip: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGCI976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 288:
SEQUENCE CHARACTERISTICS:
LENGTH: 837 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1241RP
US-08-998-416-288

[illegible]

```

RESULT 7
US-08-998-416-1137/c
; Sequence 1137, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philppsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jorgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebschule, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
; NUMBER OF SEQUENCES: 1152
; NUMBER OF INVENTION: AND USES THEREOF
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 1137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAGI692RP
US-08-998-416-1137

```

Query Match	2.4%	Score 70.2;	DB 4;	Length 636;
Best Local Similarity	47.3%;	Pred. No. 3.9e-07;		
Matches 251; Conservative	0;	Mismatches 273;	Indels 7;	Gaps 1;
Qy 1999	AACCATTTCTGTGTAGTCAATCAGTCATTAATAAGACAAGGCATCAATATTATTATAATTAT	2058		
Db.	564	AAATAATTAATTTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA		
Qy 2059	GACAAATTTGATGATTTAGATGAAAGTAGAAGATATATCTTCAATAAACCAGCTGAATAT	2118		
Db	504	TAAAAATTAATAATAAGAAATTAAGTTAAATTTAAATTTTAAATAATAATCTTCTATAA	--- 448	
Qy 2119	CAAGAACAATAATGGATATGTTAAAAAGATCTTAAACCTCATTCAAATAATAAGATGCCAAATC	2178		
Db	447	---AAAGATTAATAATAATAATAATCAACATAATAATTTATAATAATAGATATTATAATAA	392	
Qy 2179	GCAGTCAGTATTTTCTATCCCAATACATTAACAGCGCTTAGTGAAAAAACTAAACAATATT	2238		
Db	391	AATAATAATTTTACAATAATTTAAATAAATAATAATCTTTTATAATAATAAATAATATATTTT	332	


```

; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 86:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19390 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; US-08-961-527-86
;
; Query Match 2.0% Score 61; DB 4; Length 19390;
; Best Local Similarity 48.7%; Pred. No. 0.00013;
; Matches 166; Conservative 0; Mismatches 175; Indels 0; Gaps 0;
;
QY 1353 AATACCTTTAGTTTCTATTTATATCCCGCGTTATAAAGTGTGCAATTTATATTCAAAGATG 1412
   || || || || || || || || || || || || || || || || || || || || || ||
Db 4585 AACTGCATTAAATAGTGTGATTGTGCCAGTCTATAAATGTGGCGCAGTACCTAGAAAAATC 4526
   || || || || || || || || || || || || || || || || || || || || || ||
QY 1413 TGTAGATAGTCTCTTAATCAAACTGTTGTGCGATCTCGAGGTTCCTATTGTGAACGATGG 1472
   || || || || || || || || || || || || || || || || || || || || || ||
Db 4525 GATAGCTTCCCATTCAGAAAGCAGACCTATCAAAATCTGGAATATTCTTGTGTGATGAGG 4466
   || || || || || || || || || || || || || || || || || || || || || ||
QY 1473 TTCACAGATAATACCTTTAGAAAGTGATCAATAAGCTTTATGGTAATAATCTTAGGGTAGG 1532
   || || || || || || || || || || || || || || || || || || || || || ||
Db 4465 TGCAACAGATGAAAGTGGTTCGCTTGTGTGATTCAATCGCTGAACAAGATGACAGGGTGTC 4406
   || || || || || || || || || || || || || || || || || || || || || ||
QY 1533 CATCATGTCTAAACCAAAATGCGGGAATAGCCCTCAGCATCAATAGCAGCGGTTTCTTTTGC 1592
   || || || || || || || || || || || || || || || || || || || || || ||
Db 4405 AGTGGCTTCATAAAAAGAACGAAAGGATGTCCGAACGACGAAATATGGGATGAACGAGGC 4346
   || || || || || || || || || || || || || || || || || || || || || ||
QY 1593 TAAAGGTTATTACATTGTGGCAGTTAGATTTCAGATGATTATCTTTGAGCCCTGATGCGAGTTGA 1652
   || || || || || || || || || || || || || || || || || || || || || ||
Db 4345 TCACGGGATTATCTGATTTTTTATTGACTCAGATGATTATATCCATCCAGAAATGATTCA 4286
   || || || || || || || || || || || || || || || || || || || || || ||
QY 1653 ACTGTGTTTTAAAGAATTTTTTAAAAAGATAAAACGCTAGCTT 1693
   || || || || || || || || || || || || || || || || || || || || || ||
Db 4285 GAGCTTATATGAGCAATAGTTTCAAGAAAGATCGCGAATGTTT 4245
   || || || || || || || || || || || || || || || || || || || || || ||

```

```

RESULT 12
US-08-961-083-199
; Sequence 199, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
;

```

```

ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961.083

FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 199:
SEQUENCE CHARACTERISTICS:
LENGTH: 2023 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-083-199

Query Match 2.0%; Score 60.6; DB 3; Length 2023;
Best Local Similarity 44.6%; Pred. No. 8.1e-05;
Matches 237; Conservative 0; Mismatches 294; Indels 0; Gaps

Qy 1411 TGTGTAGATAGTCTCTTAATCAAACACTGTTGCGATCTCGAGGTGTTGATTTGTAACGAT 1470
Db 2 TGCCTAGATAGTATTAATTAATACTCAAACATATAAAAATATGAGATTTGTCGTTAATGAT 61

Qy 1471 GGTTCACAGATAAATACCTTAGAAGTGATCAATAAGCTTTATGGTAATAATCTCTAGGTA 1530
Db 62 GGTTCTAGCGGATTTATGTGACCTTTGTAAGAAATTTTCAGAAATGGATCACCGAAT 121

Qy 1531 CGATCATGTCTAAACCAATGCGGAATAGCCTCAGCATCAATGAGCGGTTTCTTTT 1590
Db 122 CTCATATAGAACAGAAATGCTGCTCTCTCGCCGCCAACAACCGGTCTGAATAAT 181

Qy 1591 GCTAAAGGTTTATACATTTGGCAGGTTAGATTCAGATGATTATCTTTAGCCGTGATGCAGTT 1650
Db 182 ATGTCGCGAAATTTATGTGACCTTTGTGGACTCGGATGATTGGAATGAGCAAGATTATGTA 241

Qy 1651 GAACGTGTTTAAAGAAATTTTAAAGATATAAACCGCTAGCTGTGTTTATACCACTAAT 1710
Db 242 GAAACTCTATATAAAAAATAGTAGAGTATCAGGCTGATATTGCAAGTTGGTAATTTATAT 301

Qy 1711 AGAAACGTCAATCGGATGGTAGCTTTAATCGCTAATGTTTACAATTTGCCAGAATTTTCA 1770
Db 302 TCTTTCAACGAAGTGAAGAAATGTTCTACTTTCATATATTGGGAGACTCTCTATTATGAG 361

Qy 1771 CGAAGAAACTCAACGGCTATGATGTCTACCATTTTAGAATCTTTTACGATTAGAGCT 1830
Db 362 AAGCTATATGATATGTTTCTATCTTTTGAAGACTTGTATGAAACTCAAGAAATGAAGAGT 421

Qy 1831 TGGCATTTAAGCGATGGATTTAACGAAAAATATTTGAAACGCCGCTGGATTATGACATGTC 1890
Db 422 TTTGCTTTGATATCTGCTTGGGGTAACTCTATAGGCAAGATTTCTTTGAGCAGTTGGCC 481

Qy 1891 CTAAACTCAGTGAAGTTGAAATTTTAAACATCTTAAATTAATCTGCTAT 1941
Db 482 TTTGACATAGTAAATTTAGGAGAGATGGTTACCTCAATCAAAAGGTATATA 532

```


[illegible]

```
QY 2032 AGACAGGCATCAATATATATATATATATGACAAATTTGATGATTAGATGAAAGTAGAAG 2091
      |||||
Db 3217 ATATAATAAGAGATATATATATAA--AACTAAATTTAATAATAACATCATAGGAAATG 3160
      |||||
QY 2092 TATATCTTCAATAAACCCTGAATATCAAGAGAAATGGATATGTAAAGATCTTAAA 2151
      |||||
Db 3159 GATTTGGTTTACTTTCACGCTTTATATATGAAAAA---ATACAAATTAATTTAAACAT 3103
      |||||
QY 2152 CTCATTCAAAATAAAGATGCCAAAATCGCAGTCAGTATTTCTATCCCAATACATTAAAC 2211
      |||||
Db 3102 AATATTACAAAAGATATTCTTAAATTTGAAAAATTTAAATCTTTAGTATCTTTGGAAATA 3043
      |||||
QY 2212 GCGTTAGTGAAGAACTAACAATATATTGATATATAATAAAAAATATATTCGTTATTATT 2271
      |||||
Db 3042 ATCATAAATTTAAGCTCTTTGATTTTATATAAATTTATACATTAATTTTAAATATA 2983
      |||||
QY 2272 CTCATGTGTTGATAAGAAATCATCTTTACACGAGCATCAAAAAAGAAATATTTGGTTTCTAT 2331
      |||||
Db 2982 TAAAAACTAAATTAGTAATCATCTAAATATTACAAATGAAGAAATATTATTGTAAAT 2923
      |||||
QY 2332 CATAAGCACCAAGTGAATATTTTACTAAATAATGACATCTCATATTACAGGAGTAATAGA 2391
      |||||
Db 2922 AATTATAGTAATATATTATGGAATGAATAGTAAAAAATTTAAAGAAATCATTTAAATTA 2863
      |||||
QY 2392 CTAAATAAACTGAGGCACATTTTAAGTAATATTAAATAATTAAGTCAGTTAAATCTAAAT 2451
      |||||
Db 2862 GTTTAAATATTTTAAATAAAATTAATAATTAATAAATAAACAATAAATAATTAAGTTCAAT 2803
      |||||
QY 2452 TGTGAATACATCATTTTGTGATAATCATGACAGCCTATTCGTTAAAAATGACAGCTATGCT 2511
      |||||
Db 2802 TGGGACTTACTCTATTTTATTAAAGATTTAAGAAATAGTTAATAATTAATAATTTGAAAAATAT 2743
      |||||
QY 2512 TATATGA 2518
      |||||
Db 2742 GATATAA 2736
```

```

;
; LENGTH: 811 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-083-197

Query Match
Best Local Similarity 50.7%; DB 3; Length 811;
Matches 138; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 1411 TGTGTAGATAGTCTCTTAAATCAAACTGTGTGATCTCGAGGTTTCTATTGTGAACGAT 1470
      |||||
Db 2 TGTGTTGGATAGCATTCAGAAATCAGACGTATCAAAATTTTGGAGTGTATTAAATCAATGAT 61
      |||||
QY 1471 GGTTCACAGATAATACCTTTAGAAGTGATCAATAAGCTTTATGGTAATAATCCTAGGGTA 1530
      |||||
Db 62 GGCTCTCCAGATCATTCACCAAAATATGTGAAGAATTTGTAGAGAAAGATTTCTCGTTTC 121
      |||||
QY 1531 CGCATCATGTCTTAAACCAAAATGGCGGAATAGCCTCAGCATCAAAATGACAGCCGTTCTTTT 1590
      |||||
Db 122 AATATATTTTGAGAAAGCAAAACGGCGTCTTTTCATCAGCTCGTAACCTAGGTATTGAATGT 181
      |||||
QY 1591 GCTAAAGGTTATTACATTGGGCAGTTAGATTCAGATGATTATCTTGAGCCTGATGCAGTT 1650
      |||||
Db 182 TCGGGGGGGGCGTACATTACTTTTGTAGACTCTGATGATTGGTTGGAAACATGATGCTTTA 241
      |||||
QY 1651 GAACTGTGTTTAAAGAAATTTTAAAAAGATAA 1682
      |||||
Db 242 GACCGATTATATGCTGCTTTTGAAAAAGGAAAA 273
      |||||

Search completed: January 4, 2003, 00:40:18
Job time : 314 secs
```

```
RESULT 15
US-08-961-083-197
; Sequence 197, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 197:
; SEQUENCE CHARACTERISTICS:
```


Db	181	GCACATCTTCTGTAATAAAGAGAAAAGCTCAATGTTTCGGATAGTCCGTTAGATATT	240
Qy	262	GCAACACAGCTCTTACTTTTCCACGTAATAAAATTAACCTATCCGAATCAGAAAAAAC	321
Db	241	GCAACACACTGTTACTTTTCCACGTAATAAAATTAAGTACTTCTGACTCGGAAAAAAC	300
Qy	322	AGTTTAAAAATAAATGGAATCTATCACTCGGAAAAAATCGGAGAACGACAGAAATCAGA	381
Db	301	ACGTTAAAAATAAATGGAATCTCTACTCAGAGAATAATCTGAATAATCGGAGGTAAGA	360
Qy	382	AAGTGGAACTAGTACCCAAAGATTTTCTTAAAGATCTTGTCTTCTGCTCCATTTGCCAGAT	441
Db	361	CGGTGCGCCCTGTGACCAAAAGATTTTCCCAAGATCTGGTTTATGAGCCCTTACCTGAT	420
Qy	442	CATGTTAATGATTTTACATGATACAAAATCGAAAAAAGCTTAGGTATTAAGCCTGTGA	501
Db	421	CATGTTAATGATTTTACATGATACAAAAGGAGAAAAAGACTTGGCATTAACACTGAA	480
Qy	502	AATAGATATCGGTCTTCTTATTTATTTATCTCTACATTTAATCGTAGCGTATTTTAGAT	561
Db	481	CATCAACATGTGGTCTTCTTATTTATCGTTTACAACATTTCAATTCGACACAGCAATTTTATCG	540
Qy	562	ATAAGTTAGCCTGTTTGGTCAATCAGAAAAAACTACCCATTTTGAAGTCGTTTGTCA	621
Db	541	ATTACATTAGCCTGTTTAGTAACCAAAAAACACATTACCCGTTTGAAGTTATCGTGACA	600
Qy	622	GATGATGTTAGTAAAGAAACTTACTTACCATTGTGCAAAAATACGACACAAAACCTGAC	681
Db	601	GATGATGTTAGTAAAGAAAGATCATACCGGATCATTCGCCAATATGAAATTAATTTGGAT	660
Qy	682	ATAAGTATGTAAGACAAAAGATTTATGAVATCAATTTGTGTGAGTCAGAAACTTTAGT	741
Db	661	ATTCGCTACGTCAGACAAAAGATTAACGGTTTTCAGCCAGTGTCCGCTCGGAATATGGGA	720
Qy	742	TTACGTACAGAAAAGTATGTTTGTCTCGATTCTAGACTCGGATATGGCACCAACAACAA	801
Db	721	TTACGCTTAGCAAAATATGACTTTATTGGCTTACTCGAGTGTATGCGCAAAATCCA	780
Qy	802	TTATGGGTTCACTTATCTTACAGAACTATTAGAAGCAATGATATTCTTTTAATTGGA	861
Db	781	TTATGGGTTCACTTATGTTTCAGAGCTATTAGAAGATGATGATTAAACAATCATTTGT	840
Qy	862	CCTAGAAATATGCGATCTCATATATTACCGCAGAACAAATCTCTTAACGATCCATAT	921
Db	841	CCAGAGAAATACATCGATACACACATATTGACCCAAAAGACTCTTAATACGCGAGT	900
Qy	922	TTATAGATCACTACCTGAAACCGCTACAAATAACAATCCTTCGATTTACATCAAAAGGA	981
Db	901	TTGCTGTAATCATTACCAGAGTGAACCAATAATAGTTGTCGCCAAAGGGGAAGGA	960
Qy	982	AAATATATCGTTGGATTTGGAGATTAGNACATTTCAAAAAACCGATTAATCTACGCTATGT	1041
Db	961	ACAGTTTCTCTGGAATGGCGCTTAGAACAAATTCGAAAAACAGAAATCTCCGCTTATCC	1020
Qy	1042	GATCTCCGTTTCGTTTATTTTGTGGCGGTAAATGTTGCAATTTCTAAAGAAATGGCTAAAT	1101
Db	1021	GATTCGCCTTCCGTTTTTTTGGCGCGGTAAATGTTGCTTTCCCTAAAAAATGGCTAAAT	1080
Qy	1102	AAAGTAGTTTGGTTTCGATGAAGAAATTTAATCATTTGGGGGGGCGAAGATCTAGAAATTTGGT	1161
Db	1081	AAATCCGGTTCTTTGATCAGGAATTTAATCACTGSGGTGGAGAAGATGTGGAATTTGGA	1140
Qy	1162	TACAGATTAATTTGCCAAAGCTGTTTTTTCAGAGTAATTTGACGGGGGAATGGCCATCCAT	1221
Db	1141	TATCGCTTATTCGTTTACGTTAGTCTTCTTTAAACCTATTGATGGCATTTATGGCCTACCAT	1200
Qy	1222	CAAGAACCACTGGTTAAAGAAATGAACAGACGCGAGCTGTAAGAACTATTACGCTT	1281
Db	1201	CAAGAGCCACAGGTTAAAGAAATGAACCGATTCGTGAAGCGGGGAAAAAATATTACGCTC	1260
Qy	1282	AAAATTTGTGAAGAAAAGGTACCTTACATCTATAGAAAGCTTTTACCATATAGAAGATCA	1341


```

; LOCATION: (3982)..(4953)
; NAME/KEY: CDS
; LOCATION: (5009)..(5947)
US-09-900-038A-3

```

Query Match 2.7%; Score 80.6; DB 10; Length 6865;
Best Local Similarity 53.3%; Pred. No. 8.1e-07;
Matches 170; Conservative 0; Mismatches 149; Indels 0;

Qy	1361	TAGTTCTTATTTATATATCCCGCTTATTA	CTGCGAAATATATATCAAGATGTTAGATA	1420
Db	3998	TTGTTCTTATCGTTATACCTATATCAACT	CTCGAAGCATATCTTTAAAGATGCGTGCAAT	4057
Qy	1421	GTGCTCTTAATCAAACTGTTGTCGATCTC	GCAGGTTTGTATTGTTAAAGCATGGTTCAACAG	1480
Db	4058	CCGTCTCAACACAGACTCATCTCATTGAT	GAGAAGTTATACTGATTAAATGATGGATCCCACTG	4117
Qy	1481	ATAATACCTTTAGAAAGTGATCAATAGCT	TTATGTTAATAATCTTAGGGTAGGCATCATGT	1540
Db	4118	ATAATAGTGGAGAAATTTCTGTATAATTT	ATCTCAAAAAGACGATCGCATCTGTTATTC	4177
Qy	1541	CTAAACCAATGCGGAATAGCCTCAGCATCA	AAATGSCAGCGGTTCTTTGCTTAAAGGTT	1600
Db	4178	ATAAAAAAATGAGGGGTATCTCGGCAAG	GAACCTAGGCTCTTGATAAATCCACAGCGC	4237
Qy	1601	ATTACATTTGGCGAGTTAGATTACAGATGA	TATTCATTGACGCTTGATCCAGCTTCGAATCTGTTT	1660
Db	4238	AATTCATAGCTTTGTGTAGATGATGATTGT	TGTAGCACCGGAATATAATTGAAAATAATGT	4297
Qy	1661	TAAAAAATTTTTTAAAGA	1679	
Db	4298	TAAAAAATTTAATCACTGA	4316	

RESULT 4

US-09-816-028A-1/c
; Sequence 1, Application US/09816028A
; Patent No. US20020042369A1

GENERAL INFORMATION:
 APPLICANT: Gilbert, Michel
 APPLICANT: Wakarchuk, Warren W.
 APPLICANT: National Research Council of Canada
 TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
 TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
 TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
 FILE REFERENCE: 019633-000111US
 CURRENT APPLICATION NUMBER: US/09/816,028A
 CURRENT FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: US 60/118,213
 PRIOR FILING DATE: 1999-02-01
 PRIOR APPLICATION NUMBER: US 09/495,406
 PRIOR FILING DATE: 2000-01-31
 NUMBER OF SEQ ID NOS: 49
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 1
 LENGTH: 11474
 TYPE: DNA
 ORGANISM: Campylobacter jejuni
 FEATURE:
 OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
 OTHER INFORMATION: including LOS biosynthesis locus
 US-09-816-028A-1

	Query Match	2.7%	Score 79.2	DB 10	Length 11474
	Best Local Similarity	51.2%	Pred. No. 1.7e-06		
	Matches 211	Conservative	138	Indels	3 Gaps
QY	1296	AAAGTCACCTTACATCTATAGAAAGCTTTTACC	AAATAGAAAGATTACATATTCATAGAAT	1355	
DB	11424	AAATTTCCTAACCTTGTATAGATTTGTTTTATT	TTTATTTTAA	CAAAATTTAGGAAAAATAT	11365
QY	1356	ACGCTTTAGTTCTATTATATCCCGCGCTTTAA	CAACTGTCAAAATATATATTC	CAAAAGATGTGT	1415
DB	11364	GCACCAACTTCTATCAAAATCCGCGCTTTTAA	TTCTTTCGCGATTTATCTCA	AGAGCTTT	11305

[illegible]

RESULT 5

US-09-870-759-83
; Sequence 83, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:

```

1 GENERAL INFORMATION:
2 APPLICANT: TERMAN, David S
3 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
4 FILE REFERENCE: 870759
5 CURRENT APPLICATION NUMBER: US/09/870,759
6 CURRENT FILING DATE: 2002-01-14
7 PRIOR APPLICATION NUMBER: US 60/208,128
8 PRIOR FILING DATE: 2000-05-30
9 NUMBER OF SEQ ID NOS: 166
10 SOFTWARE: PatentIn version 3.1
11 SEQ ID NO 83
12 LENGTH: 17276
13 TYPE: DNA
14 ORGANISM: Streptococcus agalactiae
15 FEATURE:
16 NAME/KEY: CDS
17 LOCATION: (7062)..(8207)
18 OTHER INFORMATION:
19 US-09-870-759-83

```

Query Match 2.5%; Score 74.2; DB 9; Length 17276;
Best Local Similarity 48.7%; Pred. No. 1.8e-05;
Matches 202; Conservative 0; Mismatches 213; Indels 0; Gaps 0;

[illegible]

[illegible]

Db 1528 AAAAATATTCAGATATTTCAAGAGGATATATATAAAAAATAAAGCCAAAAGGATAACACT 1587
Qy 2347 AATATTTTACTAAATAGCATCTCATATTTACACGAGTATACAGTAATAAAAACTGAG 2406
Db 1588 GAAATGTTAGATATATATAGGAATA-----ACAATATTTGATATAAAAAATTTGAT 1641
Qy 2407 GCACATTTAAGTAAATTAATTAAGTCAGTTAAATCTTAAATCTTGAATACATCAT 2466
Db 1642 GATATAAAAAATGTTGGAGATATAAAAAAGTTGCGAGATATAAAAAAGTTGATGATATA 1701
Qy 2467 TTTGATATCATGACACGCTATTCGTTTAAAAATCACAGCTATGCTTATATGAAAAATAT 2526
Db 1702 AACAACTTTGATGCTATATAAAAAATGTTGATGCTATAAAAAATGTTGATGTTAAAAAAT 1761
Qy 2527 GATCTGGCATGAATTTCTCAGCATTTAACACATGATTGGATCGAGAAAAATCAATGC 2582
Db 1762 GTTGATGCTATAAACAATGTGGAGATATAAACAATGCTGGAGATACAATAATGC 1817

RESULT 11
US-09-816-028A-26
; Sequence 26, Application US/09816028A
; Patent No. US20020042369A1
; GENERAL INFORMATION:
; APPLICANT: Gilibert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 906
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(906)
; OTHER INFORMATION: beta-1,3-galactosyltransferase from C. jejuni strain
; OTHER INFORMATION: OH4384 (ORF 6a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus
US-09-816-028A-26

Query Match 2.2% Score 66.6; DB 10; Length 906;
Best Local Similarity 51.3%; Pred. No. 0.00026;
Matches 181; Conservative 0; Mismatches 189; Indels 3; Gaps 1;

Qy 1362 AGTTTCTATTATATCCCGCTTATAACTGTGCAAAATATATATCAAGATGTGTAGATAG 1421
Db 9 AATTTCAATCATCTTACCAACTTATAATGTGGAACATATATATAGCAAGGCAATAGAAG 68
Qy 1422 TGCCTTAAATCAAACTGTTGTCGATCTCGAGGTTTGTATTGTAACTGATGTTCAACAGA 1481
Db 69 CTGTATCAATCAGACTTTTAAAGATATAGAAAATAATTGTAGTTGATGATTTGGAAATGA 128
Qy 1482 TAATACCTTAGAAGTGATCAATAGCTTTATGTTAATTAATCCTAGGGTACGCATCATGTC 1541
Db 129 TAATAGTATAATATAGCCAAGAATACTCTAAAAGACAAAAAATAAATAAATCA 188
Qy 1542 TAAACCAATGGC---GGAATAGCCTCAGCATCAAAATGCAAGCCGTTCTTTTGTGTAAGG 1598
Db 189 CAATGAAAAAAACCTTAGGCTCTTTAAGGACGAAGATATGAAGGTGTGAAGTAGCAACTC 248
Qy 1599 TTATTACATTTGGGAGTTAGATGATGATATCTTGTAGGCTGATGCGATTTGAAGCTGTG 1658
Db 249 TCCTTATATAATGTTTTTAGATCTCTGATGATTTATTTGGAACATAAATGCTTGTGAAGAGTG 308

Qy 1659 TTTAAAGCAATTTTAAAGATAAAACGCTAGCTGTGTGTTTATACCACTAATA 1711
Db 309 TATAAAATTTTAGATGAACAGGATGAAGTTGATTTAGTTGTTTTTCAATGCTA 361

RESULT 12
US-09-070-927A-308/c
; Sequence 308, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
; APPLICANT: Patrick J. Dillon
; APPLICANT: Steven Barash
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 382
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB369
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 308:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2406 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 308:
US-09-070-927A-308

Query Match 2.1% Score 61.8; DB 10; Length 2406;
Best Local Similarity 50.0%; Pred. No. 0.0028;
Matches 147; Conservative 1; Mismatches 146; Indels 0; Gaps 0;

Qy 1371 TTATATCCCGCTTATAACTGTGCAAAATATATATCAAGATGTGTAGATAGTCTCTTAA 1430
Db 2311 TTGTNCCNGTCTCTATTAACTGGAACCCCTATTTTGGNGGAGGCATTAATCAGTTTGA 2252
Qy 1431 TCAAACTGTTGTCGATCTCGAGGTTTGTATTGTTAACTGTTTCAACAGATTAATACCT 1490
Db 2251 CCAAACTATGAAGAATTTGAATCTTAAATTAATGATGCTCTACTGATAATAGTCA 2192
Qy 1491 AGAAGTGATCAATAGCTTTATGGTAAATATCTAGGGTACGCATCATGCTTAAACCAA 1550
Db 2191 AAAAATTTTGAAGAAATGCAAGATGATCCAGCTTTTCCCGTTTTCATATGTTAA 2132
Qy 1551 TGGCGGAATAGCCTCAGCATCAATGCAAGCCGTTCTTTTGTAAAGGTTATTACATTGG 1610
Db 1551 TTTTAAAAATTTTAAAGATAAAACGCTAGCTGTGTGTTTATACCACTAATA 1711
Db 309 TATAAAATTTTAGATGAACAGGATGAAGTTGATTTAGTTGTTTTTCAATGCTA 361

Db 2131 TGGGGGAATTGGAAGAAGCCTTTAACTTGGGTGTTTCAGNAGCAAAAGGTGAGTATATTC 2072

Qy 1611 GCAGTTAGATTGAGATGATTTCTTGGAGCCTGATGCAGTTGAACTGTGTTTAAA 1664

Db 2071 TGAATTGAAAGTACGATATATGTAGCATTCATGCGTATGAGCGTTTATACAA 2018

RESULT 13

US-09-765-272-199

; Sequence 199, Application US/09765272

; Patent No. US20020061545A1

; GENERAL INFORMATION:

; APPLICANT: Choi et al.

; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines

; NUMBER OF SEQUENCES: 452

; CORRESPONDENCE ADDRESS:

; ADDRESS: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/765,272

; FILING DATE: 22-Jan-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/961,083

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Brookes, A. Anders

; REGISTRATION NUMBER: 36,373

; REFERENCE/DOCKET NUMBER: FB340P2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 199:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2023 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 199:

US-09-765-272-199

Query Match 2.0%; Score 60.6; DB 10; Length 2023;

Best Local Similarity 44.6%; Pred. No. 0.0046;

Matches 237; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

Qy 1411 TGTGTAGATGTCCTTAATCAACATGTTGTCGATCGAGGTTTGTATTTGTAACGAT 1470

Db 2 TGCCTAGATAGTATTACTCAACATATATAAATATGAGATTGTTGTCGTTAATGAT 61

Qy 1471 GGTTCACAGATATACCTTAGAGTGATCAATAGCTTTATGGTAATATCTTAGGTA 1530

Db 62 GGTTCACGAGTCTTACGTTGAATTTGTAAAGAAATTTTCAGAAATGATCACCGAAT 121

Qy 1531 CGCATCATGTCTAAACCAATGGCGAATAGCCTCAGCATCAAAATGACCGGTTTCTTTT 1590

Db 122 CTCATATAGAACAGAAATGCTGCTCTTCTGCGCACGAAACACCGGTCGTAATAT 181

Qy 1591 GCTAAGGTTATACATGGCGAGTTAGATTCAGATGATTATCTTGACCTGATGCACT 1650

Db 182 ATGTCGGAAATATGTGACCTTTGTGACTCGGATGATGGATTGACGAGATTATGTA 241

Qy 1651 GAACGTGTTTAAAGAAATTTTAAAGATAAAACCGCTAGCTGTGTTTATACCACTAAT 1710

Db 242 GAACTCTATATAAAATAAGTACGATATCAGGCTGATATTCAGCTTGGTAATTTAT 301

Qy 1711 AGAAGCTCAATCGSATGGTAGCTTAATCGCTAATGGTTACAATGGCCAGAAATTTCA 1770

Db 302 TCTTTCAAGGAAGTGAAGGAATGTTCTACTTTCATATATTTGGAGACTCCTATTATGAG 361

Qy 1771 CGAGAAAAACTCACAACGGCTATGATTGCTCACCATTTTAGAAATGTTTAGCATAGCT 1830

Db 362 AAAGTATATGATAATGTTTCTCTCTTTGAGAATCTCTATGAAACTCAAGAAATGAAGT 421

Qy 1831 TGGCATTTAAGGATGATTTAACGAAATATTTGAACCCCGCTGGATTATGACATGTTTC 1890

Db 422 TTTGCTTTGATATCTGTTGGGGTAACTCTATAGGCAAGATTGTTTGACGAGTTCCGC 481

Qy 1891 CTTAAACTCAGTGAAGTTGGAAAAATTTAAACATCTTTAAATAAAATCTGCTAT 1941

Db 482 TTTGACATAGGTAAATTAGGAGAAGATGGTTACCTCAATCAAAAGGTATAT 532

RESULT 14

US-09-960-352-5785/C

; Sequence 5785, Application US/09960352

; Patent No. US20020137139A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Tao, Nengbing

; APPLICANT: Byatt, John C.

; APPLICANT: Mathialagan, Nagappan

; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND

; FILE REFERENCE: 16511.006/37-21(10298)C

; CURRENT APPLICATION NUMBER: US/09/960,352

; CURRENT FILING DATE: 2001-09-24

; NUMBER OF SEQ ID NOS: 15112

; SEQ ID NO 5785

; LENGTH: 516

; TYPE: DNA

; ORGANISM: Bos taurus

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (76),(90)

; OTHER INFORMATION: unsure at all n locations

; OTHER INFORMATION: Clone ID: 25-LIB3057-018-Q1-K1-G1

US-09-960-352-5785

Query Match 2.0%; Score 58.8; DB 10; Length 516;

Best Local Similarity 45.4%; Pred. No. 0.0073;

Matches 207; Conservative 0; Mismatches 249; Indels 0; Gaps 0;

Qy 2025 ATTAATAGACAAGGCATCAATTATTATAATTATGACAAATTTTGATGATTTAGATGAAAG 2084

Db 514 ATAATAAAAAATAAAAAATAATTTAAATATAAAAAAAATAATAATAATAATAAAAA 455

Qy 2085 TAGAAAGTATATCTTCATAAAAAACCGCTGAATATCAAGAAAGATGATGTTTAAAAA 2144

Db 454 AAAATAATAATAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAA 395

Qy 2145 TCTTAACTCAATCAAAATAAGATGCCAAATCCAGTCAGTATTTTCTATCCCAATAC 2204

Db 394 AAAAAAATAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAA 335

Qy 2205 ATTAACGGCTTAGTGAAAAAACTAAACAATATTATTGAATAATAATAATAATAATAT 2264

Db 334 AAAAAAATAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAA 275

Qy 2265 TATTATCTACATGTTGATAGAAATCTCTTACACGACATCAAAAAAGAAATATTGGC 2324

Db 274 AAATAAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAA 215

Qy 2325 TTTCTATCATAGCACCAAGTGAATATTTTACTAAATATGACATCTCATATTTACACGAG 2384

Db 214 APTAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAA 155

Qy 2385 TAATAGACTAATAAAAACTGAGGCACATTTAAGTAATATTAATAATTAAGTCAGTTAA 2444

Db 154 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 95
Qy 2445 TCTAAATTGTCATCATCATTTTGTGTAATCATGA 2480
Db 94 ATATNAAACTAACCCCTCNATACCTTATTGCCAATTA 59

RESULT 15
US-09-960-352-5558/c
; Sequence 5558, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Mengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 5558
; LENGTH: 431
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 24-LIB3057-024-Q1-K1-F7
US-09-960-352-5558

Query Match 2.0%; Score 58.6; DB 10; Length 431;
Best Local Similarity 47.6%; Pred. No. 0.0076;
Matches 172; Conservative 0; Mismatches 189; Indels 0; Gaps 0;
Qy 2104 AAAACCGCTGAATATCAAGAAGAAATGGATATGTTAAAGATCTTTAAACTCATTCCTCAAAAT 2163
Db 431 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 372
Qy 2164 AAGATGCCAAATCGCAGTCAGTATTTCTATCCCAATACATTTAAACGGCTTAGTGAA 2223
Db 371 TAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 312
Qy 2224 AAACATAACAATATTTGAATATAATAAATAATATATTTCTGTTATTTCTACATGTTGAT 2283
Db 311 ATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 252
Qy 2284 AAGATCATCTTACACAGACATCAAAAAGAAATATTTGGCTTTCTATCATTAAGCACC 2343
Db 251 AAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 192
Qy 2344 GTGAATATTTTACTAATAATGACATCTCATATTACACGAGTAATAGACTAATAAAACT 2403
Db 191 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 132
Qy 2404 GAGGCACATTTAAGTAATATTAATAAATAAAGTCAAGTTAAATCTAAATTTGTAATACATC 2463
Db 131 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 72
Qy 2464 A 2464
Db 71 A 71

Search completed: January 4, 2003, 02:22:36
Job time : 277 secs

THIS PAGE BLANK (USPTO)

BEST AVAILABLE COPY

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2003, 21:20:09 ; Search time 5768 Seconds
(without alignments)
12985.345 Million cell updates/sec

Title: US-09-842-484a-1
Perfect score: 2979
Sequence: 1 ttataaactgattaaagaag.....acatttgcattttattataaaa 2979

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 24791104 seqs, 12571243825 residues 49582208
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US099D_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US101A_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US101B_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US102B_COMB.seq.*

44: /cgn2_6/ptodata/1/pna/US6000_COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US6001_COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US6002_COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US6003_COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US6004_COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US6005_COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US6006_COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US6007_COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US6008_COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US6009_COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US6010_COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US6023_COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US6031_COMB.seq.*
76: /cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
77: /cgn2_6/ptodata/1/pna/US6033_COMB.seq.*
78: /cgn2_6/ptodata/1/pna/US6034_COMB.seq.*
79: /cgn2_6/ptodata/1/pna/US6035_COMB.seq.*
80: /cgn2_6/ptodata/1/pna/US6036_COMB.seq.*
81: /cgn2_6/ptodata/1/pna/US6037_COMB.seq.*
82: /cgn2_6/ptodata/1/pna/US6038_COMB.seq.*
83: /cgn2_6/ptodata/1/pna/US6039_COMB.seq.*
84: /cgn2_6/ptodata/1/pna/US6040_COMB.seq.*
85: /cgn2_6/ptodata/1/pna/US6041_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	2979	100.0	2979	1	PCT-US01-1335-1 Sequence 1, Appl
2	2979	100.0	2979	1	PCT-US99-2651-4 Sequence 4, Appl1
3	2963	99.5	2979	1	PCT-US01-1335-3 Sequence 3, Appl1
4	2285	76.7	2937	38	US-10-011-766B-8 Sequence 8, Appl1
5	2285	76.7	2937	38	US-10-011-771B-8 Sequence 8, Appl1
6	2283.4	76.6	2937	1	PCT-US99-07289-2 Sequence 2, Appl1
7	2283.4	76.6	2937	16	US-09-283-402-2 Sequence 2, Appl1
8	2283.4	76.6	2937	18	US-09-469-200-2 Sequence 2, Appl1
9	2283.4	76.6	2937	33	US-09-879-959-9 Sequence 9, Appl1
10	2283.4	76.6	2937	41	US-10-172-527-9 Sequence 9, Appl1
11	1503.2	50.5	2112	1	PCT-US99-26501-2 Sequence 2, Appl1
12	89.2	3.0	984	15	US-09-107-532-2527 Sequence 2527, Ap
13	89.2	3.0	984	15	US-09-107-532A-2527 Sequence 2527, Ap
14	86	2.9	993	15	US-09-134-000-2987 Sequence 2987, Ap
15	86	2.9	5550	14	US-09-070-927-137 Sequence 137, App
16	86	2.9	5550	14	US-09-070-927A-137 Sequence 137, App
17	82.2	2.8	908	26	US-09-666-355A-6574 Sequence 6574, Ap
18	82	2.8	1368	38	US-10-015-127-2421 Sequence 2421, Ap
19	80.6	2.7	6865	34	US-09-900-038A-3 Sequence 3, Appl1
20	79.4	2.7	1095	38	US-10-015-127-1938 Sequence 1938, Ap
21	79.2	2.7	11474	18	US-09-495-406-1 Sequence 1, Appl1

c 22 79.2 2.7 11474 31 US-09-816-028-1 Sequence 1, Appli
c 23 79.2 2.7 11474 31 US-09-816-028A-1 Sequence 1, Appli
c 24 78.8 2.6 810 26 US-09-666-355A-6400 Sequence 6400, Ap
25 78.8 2.6 1002 15 US-09-107-532-1469 Sequence 1469, Ap
26 78.8 2.6 1002 15 US-09-107-532A-1469 Sequence 1469, Ap
27 77.6 2.6 577 33 US-09-874-708A-11415 Sequence 11415, A
28 77.6 2.6 577 65 US-60-211-750-11272 Sequence 11272, A
c 29 76.6 2.6 13825 24 US-09-634-238-29 Sequence 29, Appl
c 30 76.2 2.6 1141 31 US-09-806-708A-22 Sequence 22, Appl
31 76 2.6 7458 1 PCT-US99-26796-239 Sequence 239, App
32 75.8 2.5 777 17 US-09-397-761A-2730 Sequence 2730, Ap
33 75.8 2.5 777 32 US-09-838-601-2730 Sequence 2730, Ap
c 34 75.8 2.5 980 26 US-09-666-355A-6399 Sequence 6399, Ap
35 75.8 2.5 1776 38 US-10-015-127-2386 Sequence 2386, Ap
36 75.2 2.5 824 64 US-60-207-458-18907 Sequence 18907, A
c 37 75.2 2.5 1250 29 US-09-739-449-1859 Sequence 1859, Ap
c 38 75.2 2.5 1250 31 US-09-803-110-1859 Sequence 1859, Ap
39 75.2 2.5 18585 50 US-60-068-175-605 Sequence 605, App
40 75 2.5 581 38 US-10-020-338-12117 Sequence 12117, A
41 75 2.5 961 56 US-60-126-265-3039 Sequence 3039, Ap
42 74.8 2.5 948 18 US-09-404-520-13069 Sequence 13069, A
c 43 74.8 2.5 4985 39 US-10-056-405-10 Sequence 10, Appl
c 44 74.8 2.5 4985 39 US-10-094-240-10 Sequence 10, Appl
45 74.2 2.5 17276 33 US-09-870-759-83 Sequence 83, Appl

ALIGNMENTS

RESULT 1
PCT-US01-13395-1
; Sequence 1, Application PC/TUS0113395
; GENERAL INFORMATION:
; APPLICANT: DE ANGELIS, PAUL L.
; TITLE OF INVENTION: CHONDROITIN SYNTHASE GENE AND METHODS OF MAKING AND
; FILE OF INVENTION: USING SAME
; FILE REFERENCE: 618755-9/JP/199,538
; CURRENT APPLICATION NUMBER: PCT/US01/13395
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 09/437,277
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2979
; TYPE: DNA
; ORGANISM: Pasteurella multocida
PCT-US01-13395-1

Query Match 100.0%; Score 2979; DB 1; Length 2979;
Best Local Similarity 100.0%; Pred No. 0;
Matches 2979; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTATAACTGATTAAAGAGGTAACGATTCACAGCAAGGTTAAATTTTAAAGAAAGAAA 60
DB 1 TTATAACTGATTAAAGAGGTAACGATTCACAGCAAGGTTAAATTTTAAAGAAAGAAA 60
QY 61 ATGAATACATTATCAGACAGCAATAAAGCATATAACAGCAATGACTATGATTAAGCACTC 120
DB 61 ATGAATACATTATCAGACAGCAATAAAGCATATAACAGCAATGACTATGATTAAGCACTC 120
QY 121 AAATATTATTGAGAGTCTGCTGAAACCTACGGCGGCAAAATTCGTTGAATTCCTAAATATC 180
DB 121 AAATATTATTGAGAGTCTGCTGAAACCTACGGCGGCAAAATTCGTTGAATTCCTAAATATC 180
QY 181 AAATGTAAGAAAGAACTCTCGACCAATTCATTATGTAAGTGAAGATAAAAAACAGTGTT 240
DB 181 AAATGTAAGAAAGAACTCTCGACCAATTCATTATGTAAGTGAAGATAAAAAACAGTGTT 240
QY 241 TGGGATAGCTCATTAGATATCGCAACACAGCTCTTCTTCCACAGCTAAAGAAATTAAC 300

Db 241 TCGGATAGCTCATTAGATATCGCAACACAGCTCTTACTTTCCAAACGTAAAGAAATTAAC 300
QY 301 CTATCCGAATCAGAAAAACAGTTTAAAAATAAATGAAATCTATCACTGGGAAAAAA 360
Db 301 CTATCCGAATCAGAAAAACAGTTTAAAAATAAATGAAATCTATCACTGGGAAAAAA 360
QY 361 TCGGAGAACGCGAAGAAATCAGAAAGGTGGAACCTAGTACCACAAAGATTTTCTTAAGATCTT 420
Db 361 TCGGAGAACGCGAAGAAATCAGAAAGGTGGAACCTAGTACCACAAAGATTTTCTTAAGATCTT 420
QY 421 GTTCTTGCTCCATTGCGAGATCATGTAAATGATTTTACATGGTACAAAAATCGAAAAAA 480
Db 421 GTTCTTGCTCCATTGCGAGATCATGTAAATGATTTTACATGGTACAAAAATCGAAAAAA 480
QY 481 AGCTTAGGTATTAAGCCTGTAAATAAGATATCGGCTTTCTTCTATTATTAATTCCTACATTT 540
Db 481 AGCTTAGGTATTAAGCCTGTAAATAAGATATCGGCTTTCTTCTATTATTAATTCCTACATTT 540
QY 541 AATCGTAGCGGTATTTTAGATATAACGTTAGCCTGTTTGGTCAATCAGAAAAACAACTAC 600
Db 541 AATCGTAGCGGTATTTTAGATATAACGTTAGCCTGTTTGGTCAATCAGAAAAACAACTAC 600
QY 601 CCATTTGAAGTCGTTGTCGAGATGATGGTAGAAGGAAAACTTACTTACCATTGTGCAA 660
Db 601 CCATTTGAAGTCGTTGTCGAGATGATGGTAGAAGGAAAACTTACTTACCATTGTGCAA 660
QY 661 AAATACGAAACAAAACCTTGACATAAAGTATGTAAAGACAAAAAGATTATGATATCAATTG 720
Db 661 AAATACGAAACAAAACCTTGACATAAAGTATGTAAAGACAAAAAGATTATGATATCAATTG 720
QY 721 TGTGAGTCAGAACTTAGTTTACGTACAGCAAGCTATGATTTGCTCTCGATTCTAGAC 780
Db 721 TGTGAGTCAGAACTTAGTTTACGTACAGCAAGCTATGATTTGCTCTCGATTCTAGAC 780
QY 781 TGGCATATGCGACCAACAACTTATGGTTCTATTCTTATCTTACAGAACTATTAGAAGAC 840
Db 781 TGGCATATGCGACCAACAACTTATGGTTCTATTCTTATCTTACAGAACTATTAGAAGAC 840
QY 841 AATGATATTTTAAATGGACCTAGAAAAATATGTTGATCTCATATATTAATACCGCAGAA 900
Db 841 AATGATATTTTAAATGGACCTAGAAAAATATGTTGATCTCATATAATATACCGCAGAA 900
QY 901 CAATTCCTTAACGATCCATATTAATAGAACTACTACTGAAACCGCTACAAATAACAAT 960
Db 901 CAATTCCTTAACGATCCATATTAATAGAACTACTACTGAAACCGCTACAAATAACAAT 960
QY 961 CCTTCGATTACATCAAAAGGAAATATATCGTTGGATTGGAGATTAGAACAATTTCAAAAA 1020
Db 961 CCTTCGATTACATCAAAAGGAAATATATCGTTGGATTGGAGATTAGAACAATTTCAAAAA 1020
QY 1021 ACCGATATCTACGCTATGATGATTCCTCGTTTCGTTATTTTGTTCGGGTAAATGTTCCA 1080
Db 1021 ACCGATATCTACGCTATGATGATTCCTCGTTTCGTTATTTTGTTCGGGTAAATGTTCCA 1080
QY 1081 TTTTCTAAAGAAATGGCTAAATAAAGTAGTTGGTTTCGATGAAGAAATTAATCAATTTGGGG 1140
Db 1081 TTTTCTAAAGAAATGGCTAAATAAAGTAGTTGGTTTCGATGAAGAAATTAATCAATTTGGGG 1140
QY 1141 GCGGAAGATGTAGAAATTTGGTTACAGATTAATTTCCAAAAGCTGTTTTTCAGAGTAAT 1200
Db 1141 GCGGAAGATGTAGAAATTTGGTTACAGATTAATTTCCAAAAGCTGTTTTTCAGAGTAAT 1200
QY 1201 GACGGCGGAATGGCCATCCATCAAGAACCCACTGGTAAAGAAATGAACAGACGCCAA 1260
Db 1201 GACGGCGGAATGGCCATCCATCAAGAACCCACTGGTAAAGAAATGAACAGACGCCAA 1260
QY 1261 GCTGTTAAAGTATTAAGCTTAAATTTGTAAGAAAAAGGTACCTTACATCTATAGAAG 1320
Db 1261 GCTGTTAAAGTATTAAGCTTAAATTTGTAAGAAAAAGGTACCTTACATCTATAGAAG 1320
QY 1321 CTTTTCACCAATFAGAAATTCACATTTTACATAGATACCTTTAGTTTCTTATTATATCCCC 1380
Db 1321 CTTTTCACCAATFAGAAATTCACATTTTACATAGATACCTTTAGTTTCTTATTATATCCCC 1380

Qy 1381 GCTTATAACTGTGCAAAATATATCAAAAGATGTAGATAGTCTCTTAATCAAACTGTT 1440
Db 1381 GCTTATAACTGTGCAAAATATATCAAAAGATGTAGATAGTCTCTTAATCAAACTGTT 1440
Qy 1441 GTCGATCTCGAGGTTTGTATTTGTAACAGATGGTTTCAACAGATAATACCTTTAGAGTAGTC 1500
Db 1441 GTCGATCTCGAGGTTTGTATTTGTAACAGATGGTTTCAACAGATAATACCTTTAGAGTAGTC 1500
Qy 1501 AATAAGCTTTATGCTAATAATCTTAGGTCAGGCATCATGTCTAACCACCAATGGCGGATA 1560
Db 1501 AATAAGCTTTATGCTAATAATCTTAGGTCAGGCATCATGTCTAACCACCAATGGCGGATA 1560
Qy 1561 GCCTCAGCATCAAAATGACGCGGTTCTTTTGCCTAAAGGTTATACATTTGGGCGAGTTAGAT 1620
Db 1561 GCCTCAGCATCAAAATGACGCGGTTCTTTTGCCTAAAGGTTATACATTTGGGCGAGTTAGAT 1620
Qy 1621 TCAGATGATTATCTTGAGCGCTGATGCACTGAGTGTGTTTAAAGAAATTTTTTAAAGAT 1680
Db 1621 TCAGATGATTATCTTGAGCGCTGATGCACTGAGTGTGTTTAAAGAAATTTTTTAAAGAT 1680
Qy 1681 AAAACGCTAGCTGTGTTTATACCACTAATAGAAACGTCATCCGGATGGTAGCTTAATC 1740
Db 1681 AAAACGCTAGCTGTGTTTATACCACTAATAGAAACGTCATCCGGATGGTAGCTTAATC 1740
Qy 1741 GCTAATGTTTACAAATTTGCCAGAAATTTTACGAGAAAAAATCAACACGCGCTATGATGCT 1800
Db 1741 GCTAATGTTTACAAATTTGCCAGAAATTTTACGAGAAAAAATCAACACGCGCTATGATGCT 1800
Qy 1801 CACCATTTTGAATGTTTACGATTTAGAGCTTTGGCATTTTAAACGATGATTTTAAAGAAAT 1860
Db 1801 CACCATTTTGAATGTTTACGATTTAGAGCTTTGGCATTTTAAACGATGATTTTAAAGAAAT 1860
Qy 1861 ATTCAAAACCGCGGATTTAGCATCTCTTAACTCAGTGAAGTTGGAAATTTTAA 1920
Db 1861 ATTCAAAACCGCGGATTTAGCATCTCTTAACTCAGTGAAGTTGGAAATTTTAA 1920
Qy 1921 CATCTTAATAAATCTGCTATAACCGGCTATTACATGGTGATAACACATCCATTAAGAAA 1980
Db 1921 CATCTTAATAAATCTGCTATAACCGGCTATTACATGGTGATAACACATCCATTAAGAAA 1980
Qy 1981 CTCGGCATTTCAAAAGAAAAACCATTTTGTGTAGTCAATCAGTCAATTAAGTCAAGGC 2040
Db 1981 CTCGGCATTTCAAAAGAAAAACCATTTTGTGTAGTCAATCAGTCAATTAAGTCAAGGC 2040
Qy 2041 ATCAATTTATTAATATGACAAATTTGATGATTTAGTGAAGTGAAGTATATCTTC 2100
Db 2041 ATCAATTTATTAATATGACAAATTTGATGATTTAGTGAAGTGAAGTATATCTTC 2100
Qy 2101 AATAAACCCTGATATCAAGAGAAATGGATATGTTTAAAGATCTTAAACTCATTC 2160
Db 2101 AATAAACCCTGATATCAAGAGAAATGGATATGTTTAAAGATCTTAAACTCATTC 2160
Qy 2161 AATAAGATGCAAAATCGCAGTCAGTATTTCTATCCCAATACATTAACGGCTTAGTG 2220
Db 2161 AATAAGATGCAAAATCGCAGTCAGTATTTCTATCCCAATACATTAACGGCTTAGTG 2220
Qy 2221 AAAAACTAACAATATTTATGAATAATAAATAATATTTCTGTTATTTCTACATGTT 2280
Db 2221 AAAAACTAACAATATTTATGAATAATAAATAATATTTCTGTTATTTCTACATGTT 2280
Qy 2281 GATAAGATCATCTTACACAGACATCAAAAGAAATATTTGGCTTTCTATCATTAAGCAC 2340
Db 2281 GATAAGATCATCTTACACAGACATCAAAAGAAATATTTGGCTTTCTATCATTAAGCAC 2340
Qy 2341 CAAGTGAATATTTTACTAAATTAATGACATCTCATATTACAGAGTAATAGACTAATAAAA 2400
Db 2341 CAAGTGAATATTTTACTAAATTAATGACATCTCATATTACAGAGTAATAGACTAATAAAA 2400
Qy 2401 ACTGAGGCACATTTAAGTAATATTAATAATTAAGTCAAGTAAATCTAAATTTGTAATAC 2460
Db 2401 ACTGAGGCACATTTAAGTAATATTAATAATTAAGTCAAGTAAATCTAAATTTGTAATAC 2460

Qy 2461 ATCAATTTTGTATAATCATGACAGCCCTATTTCGTTAAAAATGACAGCTATGCTTTATATGAA 2520
Db 2461 ATCAATTTTGTATAATCATGACAGCCCTATTTCGTTAAAAATGACAGCTATGCTTTATATGAA 2520
Qy 2521 AAATATGATGTCGGCATGAATTTCTCAGCATTTAACACATGATGGATCGAGAAAAATCAAT 2580
Db 2521 AAATATGATGTCGGCATGAATTTCTCAGCATTTAACACATGATGGATCGAGAAAAATCAAT 2580
Qy 2581 CGCATCCACCATTTTAAAAAGCTGATTAACACCTATTTTAAATGACAAATGACTTAAGAAGT 2640
Db 2581 CGCATCCACCATTTTAAAAAGCTGATTAACACCTATTTTAAATGACAAATGACTTAAGAAGT 2640
Qy 2641 ATGAATGTGAAGGGGCATCAACAGCTATGTTTATGAAGTATGCGCTACCGCATGAGCTT 2700
Db 2641 ATGAATGTGAAGGGGCATCAACAGCTATGTTTATGAAGTATGCGCTACCGCATGAGCTT 2700
Qy 2701 CTGACGATTTTAAAGAAGTCATCAGCTCTGCCAATCAATTTGATAGTGTGCCAGAAATAT 2760
Db 2701 CTGACGATTTTAAAGAAGTCATCAGCTCTGCCAATCAATTTGATAGTGTGCCAGAAATAT 2760
Qy 2761 AACACTGAGGATATTTGGTTCCAAATTTGCACCTTTTAACTTAGAAAAAGAAACCGCCAT 2820
Db 2761 AACACTGAGGATATTTGGTTCCAAATTTGCACCTTTTAACTTAGAAAAAGAAACCGCCAT 2820
Qy 2821 GTATTTAATAAACATCGACCCCTGACTTATATGCTTTGGGAACGAAAAATACAAATGGACA 2880
Db 2821 GTATTTAATAAACATCGACCCCTGACTTATATGCTTTGGGAACGAAAAATACAAATGGACA 2880
Qy 2881 AATGAACAAATTTCAAGTGTCAAAAAAGCGGAAATATATCCCGTTTAAACGTTCAATAT 2940
Db 2881 AATGAACAAATTTCAAGTGTCAAAAAAGCGGAAATATATCCCGTTTAAACGTTCAATAT 2940
Qy 2941 AATAGTATAACGCTATAAACCAATTTTGCATTTTATTAATA 2979
Db 2941 AATAGTATAACGCTATAAACCAATTTTGCATTTTATTAATA 2979

RESULT 2

PCT-US99-26501-4

; Sequence 4: Application PC/TUS9926501

; GENERAL INFORMATION:

; APPLICANT: THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA

; TITLE OF INVENTION: POLYMER GRAFTING BY POLYSACCHARIDE SYNTHASES

; FILE REFERENCE: 617481-5

; CURRENT APPLICATION NUMBER: PCT/US99/26501

; EARLIER FILING DATE: 1999-11-10

; EARLIER FILING DATE: 1998-11-11

; EARLIER FILING DATE: 1999-04-01

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 4

; LENGTH: 2979

; TYPE: DNA

; ORGANISM: Pasteurella multocida

PCT-US99-26501-4

Query Match

Best Local Similarity 100.0%; Score 2979; DB 1: Length 2979;

Matches 2979; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTATAACTGATTAAGAAGGTAACAGTTCAAGCAAGGTTAATTTTAAAGCAAGAAA 60
Db 1 TTATAACTGATTAAGAAGGTAACAGTTCAAGCAAGGTTAATTTTAAAGCAAGAAA 60
Qy 61 ATGAATACATTATCAACAGCAATAAAGCATATAACAGCAATGACTATGAATTAGCACTC 120
Db 61 ATGAATACATTATCAACAGCAATAAAGCATATAACAGCAATGACTATGAATTAGCACTC 120
Qy 121 AAATTTTGTAGAAGTCTGCTGAAACCTACGCGGCAAAATCGTTGAATTTCAAAATATTC 180
Db 121 AAATTTTGTAGAAGTCTGCTGAAACCTACGCGGCAAAATCGTTGAATTTCAAAATATTC 180

181 AAATGTAAGAAAACTCTCGACCAATCTTATGTAAAGTAAGATAAAAAACAGTGTT 240
181 AAATGTAAGAAAACTCTCGACCAATCTTATGTAAAGTAAGATAAAAAACAGTGTT 240
241 TCGGATAGCTCATAGATATCGACACACAGCTCTACTTCCAAACGTAAGAAATTAAC 300
241 TCGGATAGCTCATAGATATCGACACACAGCTCTACTTCCAAACGTAAGAAATTAAC 300
301 CTATCCGAATCAGAAAAACAGTTTAAAAAAATAATGGAATCTATCACTGGGAAAAA 360
301 CTATCCGAATCAGAAAAACAGTTTAAAAAAATAATGGAATCTATCACTGGGAAAAA 360
361 TCGGAGAACGAGAAATCAGAAAGGTGGAAGTACAGTACCAAGATTTCTTAAAGATCTT 420
361 TCGGAGAACGAGAAATCAGAAAGGTGGAAGTACAGTACCAAGATTTCTTAAAGATCTT 420
421 GTTCTTGCCTCCATTGGCAGATCATGTTAATGATTTTACATGTTGACAAAAATCGAAAAA 480
421 GTTCTTGCCTCCATTGGCAGATCATGTTAATGATTTTACATGTTGACAAAAATCGAAAAA 480
481 AGCTTAGGTATAAGCCTGTAAATAGAAATATCGGCTCTTCTATATTTATTCCTACATTT 540
481 AGCTTAGGTATAAGCCTGTAAATAGAAATATCGGCTCTTCTATATTTATTCCTACATTT 540
541 AATCGTAGCGGTATTTAGATATAACGTTAGCCTGTTTGGTCAATCAGAAAAACAACTAC 600
541 AATCGTAGCGGTATTTAGATATAACGTTAGCCTGTTTGGTCAATCAGAAAAACAACTAC 600
601 CCATTTGAAGTCGTTGTCAGATGATGGTAAAGGAAAACTTACTACCATTTGTGCAA 660
601 CCATTTGAAGTCGTTGTCAGATGATGGTAAAGGAAAACTTACTACCATTTGTGCAA 660
661 AAATAGCAAAAACTTGACATAAGATATGTAAAGCAAAAAAGATTATGGATATCAATTC 720
661 AAATAGCAAAAACTTGACATAAGATATGTAAAGCAAAAAAGATTATGGATATCAATTC 720
721 TGTGCACTCAGAACTTAGTGTACGTACAGCAAAAGTATGATTTGCTCGATTTAGAC 780
721 TGTGCACTCAGAACTTAGTGTACGTACAGCAAAAGTATGATTTGCTCGATTTAGAC 780
781 TGGGATATGCGACACCAACAAATATGGGTTCACTTTATCTTACAGAACTATTAGAAAG 840
781 TGGGATATGCGACACCAACAAATATGGGTTCACTTTATCTTACAGAACTATTAGAAAG 840
841 AATGATATCTTTAATTTGACCTTAGAAAAATATGTGGATCTCATATATTACCGCGAAA 900
841 AATGATATCTTTAATTTGACCTTAGAAAAATATGTGGATCTCATATATTACCGCGAAA 900
901 CAATTCCTTAACGATCCATATTAATAGAACTACTACTGAAACCGCTACAAATAACAA 960
901 CAATTCCTTAACGATCCATATTAATAGAACTACTACTGAAACCGCTACAAATAACAA 960
961 CTTTCGATTACATCAAAAGAAATATATCGTTGGATTGGAGATTAGAACATTTCAAAAA 1020
961 CTTTCGATTACATCAAAAGAAATATATCGTTGGATTGGAGATTAGAACATTTCAAAAA 1020
1021 ACCGATATCTACGCTATGTGATCTCCGTTTCGTTATTTGTTGGGGTAATGTTGCA 1080
1021 ACCGATATCTACGCTATGTGATCTCCGTTTCGTTATTTGTTGGGGTAATGTTGCA 1080
1081 TTTTCTAAAGAAATGGCTAAATAAGTAGTGGTTGCTGATGAAGAAATTTAATCATTTGGGG 1140
1081 TTTTCTAAAGAAATGGCTAAATAAGTAGTGGTTGCTGATGAAGAAATTTAATCATTTGGGG 1140
1141 GCGGAAGATGTAGAAATTTGGTTACAGATATTTGCCAAAGGCTGTTTTTCAGAGTAAT 1200
1141 GCGGAAGATGTAGAAATTTGGTTACAGATATTTGCCAAAGGCTGTTTTTCAGAGTAAT 1200
1201 GACGCGGAATGGCCATCCATCAAGAACCCCTGGTAAAGAAATGAAACAGAACCGCAA 1260
1201 GACGCGGAATGGCCATCCATCAAGAACCCCTGGTAAAGAAATGAAACAGAACCGCAA 1260

1261 GCTGGTAAAGTATTACGCTTAAAAATTGTGAAGAAAAAGTACTTACATCTATAGAAG 1320
1261 GCTGGTAAAGTATTACGCTTAAAAATTGTGAAGAAAAAGTACTTACATCTATAGAAG 1320
1321 CTTTACCAATAGAAGATTCACATATTCATAGAAATACCTTTAGTTTCTATTTATATCCCC 1380
1321 CTTTACCAATAGAAGATTCACATATTCATAGAAATACCTTTAGTTTCTATTTATATCCCC 1380
1381 GCTTATAACTGTGCAAAATATATTCAAAGATGTAGATAGTCTCTTAAATCAAACTGTT 1440
1381 GCTTATAACTGTGCAAAATATATTCAAAGATGTAGATAGTCTCTTAAATCAAACTGTT 1440
1441 GTGATCTCGAGTTTGTATTTCTAAAGATGGTTCAACAGATAATACCTTTAGAGTGATC 1500
1441 GTGATCTCGAGTTTGTATTTCTAAAGATGGTTCAACAGATAATACCTTTAGAGTGATC 1500
1501 AATAAGCTTTATCGTAAATATCTAGGTACGATCATGTCTAAACCAATGCGGGAATA 1560
1501 AATAAGCTTTATCGTAAATATCTAGGTACGATCATGTCTAAACCAATGCGGGAATA 1560
1561 GCCTCAGCATCAAAATGCAGCGTTTCTTTTGTCTAAAGGTTATTACATTTGGCAGTTAGAT 1620
1561 GCCTCAGCATCAAAATGCAGCGTTTCTTTTGTCTAAAGGTTATTACATTTGGCAGTTAGAT 1620
1621 TCAGATGATTTATCTTGAGCCTGATGCGAGTTGAACTGTGTTTAAAGAAATTTTAAAGAT 1680
1621 TCAGATGATTTATCTTGAGCCTGATGCGAGTTGAACTGTGTTTAAAGAAATTTTAAAGAT 1680
1681 AAACGCTACGTTGTGTTTATACCACTAATAGAAACGCTCAATCCGATGTTAGCTTAATC 1740
1681 AAACGCTACGTTGTGTTTATACCACTAATAGAAACGCTCAATCCGATGTTAGCTTAATC 1740
1741 GCTAATGTTTACAAATTTGCCAGAAATTTTACAGAGAAAAAATCAACAGGCTATGATTGCT 1800
1741 GCTAATGTTTACAAATTTGCCAGAAATTTTACAGAGAAAAAATCAACAGGCTATGATTGCT 1800
1801 CACATTTTGAATGTTTACGATTAGAGCTTGGCATTTAACGGATGAGTTTAAACGAAAT 1860
1801 CACATTTTGAATGTTTACGATTAGAGCTTGGCATTTAACGGATGAGTTTAAACGAAAT 1860
1861 ATTGAAACGCGTGGATTTATGACATGTTCTTAAACTCAGTGAAGTTGGAAAAATTTAAA 1920
1861 ATTGAAACGCGTGGATTTATGACATGTTCTTAAACTCAGTGAAGTTGGAAAAATTTAAA 1920
1921 CATCTTAAATAAATCTGCTATAACCGGCTATTACATGGTGATAACACATCTCAATTAAGAA 1980
1921 CATCTTAAATAAATCTGCTATAACCGGCTATTACATGGTGATAACACATCTCAATTAAGAA 1980
1981 CTCGCGATTTCAAAAGAAAAACCATTTTGTGTAGTCAATCAGTCAATTAATAGACAAGGC 2040
1981 CTCGCGATTTCAAAAGAAAAACCATTTTGTGTAGTCAATCAGTCAATTAATAGACAAGGC 2040
2041 ATCAATTTATTAATTTATGCAAAATTTGATGATTTAGATGAAGTAGAAGTATATCTTC 2100
2041 ATCAATTTATTAATTTATGCAAAATTTGATGATTTAGATGAAGTAGAAGTATATCTTC 2100
2101 AATAAACCGCTGAATATCAAGAGAAATGATATGTTTAAAGATCTTAAACTCATTCAA 2160
2101 AATAAACCGCTGAATATCAAGAGAAATGATATGTTTAAAGATCTTAAACTCATTCAA 2160
2161 AATAAGATGCAAAATCGCAGTCAATTTTCTATCCCAATACATTAACGCGCTTAGTG 2220
2161 AATAAGATGCAAAATCGCAGTCAATTTTCTATCCCAATACATTAACGCGCTTAGTG 2220
2221 AAAAACTTAAACAATTTATTTGAATATATAAAAAATATATTCGTTTATTTCTACATGTT 2280
2221 AAAAACTTAAACAATTTATTTGAATATATAAAAAATATATTCGTTTATTTCTACATGTT 2280
2281 GATAGAATCATCTTACACGACACATCAAAAGAAATTTGCGTTCTCTATCATAGCAC 2340
2281 GATAGAATCATCTTACACGACACATCAAAAGAAATTTGCGTTCTCTATCATAGCAC 2340
2341 CAAGTGAATATTTTACTAAATAATGACATCTCATATTACACGAGTAATAGACTAATAA 2400

Db	2341	CAAGTGAATATTTTACCTAAATAATGACATCTCATATTACAGGAGTAATAGACTTAATAAAA	2400
Qy	2401	ACTGAGGCACATTTAAAGTAAATATTTAAATTAAGTCAGTTTAAATCTAAATTTGGAATAC	2460
Db	2401	ACTGAGGCACATTTAAAGTAAATATTTAAATAAATTAAGTCAGTTTAAATCTAAATTTGGAATAC	2460
Qy	2461	ATCATTTTTCGATATCATGACAGCGCTATTCGTTAAAAATGACACGCTATGCTTATATGAAA	2520
Db	2461	ATCATTTTTCGATATCATGACAGCGCTATTCGTTAAAAATGACACGCTATGCTTATATGAAA	2520
Qy	2521	AAATATGATGTCGGCATGAATTTCTTCAGCATTAACACATGATTCGGATCGAGAAAATCAAT	2580
Db	2521	AAATATGATGTCGGCATGAATTTCTTCAGCATTAACACATGATTCGGATCGAGAAAATCAAT	2580
Qy	2581	CGCGATCCACCATTTAAAGAGCTGATTTAAACCTATTTTAAATGACATGAGCTTAAGAAGT	2640
Db	2581	CGCGATCCACCATTTAAAGAGCTGATTTAAACCTATTTTAAATGACATGAGCTTAAGAAGT	2640
Qy	2641	ATGAATGTGAAGGGGCATCACAAAGTAGTGTATGAAGTAGTCGCCTACCCGCATGAGCTT	2700
Db	2641	ATGAATGTGAAGGGGCATCACAAAGTAGTGTATGAAGTAGTCGCCTACCCGCATGAGCTT	2700
Qy	2701	CTGACGATTTATTAAGAAGTCATCACATCTCGCCAAATCAATGTAGTGTGCCAGAAATAT	2760
Db	2701	CTGACGATTTATTAAGAAGTCATCACATCTCGCCAAATCAATGTAGTGTGCCAGAAATAT	2760
Qy	2761	AACACTGAGGATATTTGGTTCCTCAATTTGCACTTTTAACTTTAGAAAAGAAACCGGCCAT	2820
Db	2761	AACACTGAGGATATTTGGTTCCTCAATTTGCACTTTTAACTTTAGAAAAGAAACCGGCCAT	2820
Qy	2821	GTATTTTAATAACATCGACCCCTGACCTTATATGCTTTGGGAACGAAAATTTACAATGGACA	2880
Db	2821	GTATTTTAATAACATCGACCCCTGACCTTATATGCTTTGGGAACGAAAATTTACAATGGACA	2880
Qy	2881	AATCAACAAATTTAAAGTCAAAAAAAGGGCGAAATATCCCCGTTTAAACAGTTCATTATT	2940
Db	2881	AATCAACAAATTTAAAGTCAAAAAAAGGGCGAAATATCCCCGTTTAAACAGTTCATTATT	2940
Qy	2941	AATAGTAAAGCGCTATAAACAATTTGCAATTTTATTAATAA	2979
Db	2941	AATAGTAAAGCGCTATAAACAATTTGCAATTTTATTAATAA	2979

RESULT 3

```

RESULT 3
PCT-US01-13395-3
; Sequence 3, Application PC/TUS0113395
; GENERAL INFORMATION:
; APPLICANT: DE ANGELIS, PAUL L.
; TITLE OF INVENTION: CHONDROITIN SYNTHASE GENE AND METHODS OF MAKING AND
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 618755-9/JRP/199,538
; CURRENT APPLICATION NUMBER: PCT/US01/13395
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 09/437,277
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2979
; TYPE: DNA
; ORGANISM: Pasteurella multocida
PCT-US01-13395-3

```

Query Match	99.5%	Score 2963;	DB 1;	Length 2979;
Best Local Similarity	99.7%	Pred. No. 0;		
Matches 2969;	Conservative	0;	Mismatches 10;	Indels 0;
Gaps	0;			

QY 1141 GCGAAGATGTAGAATTGGTTACAGATTATTTGCCAAGGCGTGTCTTTTTCACAGTAATT 1200
DB 1141 GCGAAGATGTAGAATTGGTTACAGATTATTTGCCAAGGCGTGTCTTTTTCACAGTAATT 1200
QY 1201 GACGGCGGAATGCCATCCATCAAGAACCACCTGGTTAAAGAAAATGAACAGAACCGGAA 1260
DB 1201 GACGGCGGAATGCCATCCATCAAGAACCACCTGGTTAAAGAAAATGAACAGAACCGGAA 1260
QY 1261 GCTGGTAAAGTATACGCTTAAATTTGTAAGAAAAGGTACCTTACATCTATAGAAAG 1320
DB 1261 GCTGGTAAAGTATACGCTTAAATTTGTAAGAAAAGGTACCTTACATCTATAGAAAG 1320
QY 1321 CTTTAAACCAATGAAGATTCACATATTCATAGAAATACCTTTAGTTCTCTATTTATATCCCC 1380
DB 1321 CTTTAAACCAATGAAGATTCACATATTCATAGAAATACCTTTAGTTCTCTATTTATATCCCC 1380
QY 1381 GCTTATAACTGTGCAAAATATATTCAAAGATGTAGATAGTGTCTTAAATCAAACTGTT 1440
DB 1381 GCTTATAACTGTGCAAAATATATTCAAAGATGTAGATAGTGTCTTAAATCAAACTGTT 1440
QY 1441 GTCGATCTCGAGGTTGTATTTGTAACGATGTTCAACAGATAATACCTTAGAAGTGATC 1500
DB 1441 GTCGATCTCGAGGTTGTATTTGTAACGATGTTCAACAGATAATACCTTAGAAGTGATC 1500
QY 1501 AATAAGCTTTATGGTAAATATCTAGGTACGATCATGTCTAAACCAATGGCGGAATA 1560
DB 1501 AATAAGCTTTATGGTAAATATCTAGGTACGATCATGTCTAAACCAATGGCGGAATA 1560
QY 1561 GCCTCAGATCAAAATGCAGCGGTTCTTTTGCCTAAAGGTTATTTACATTTGGCGAGTTAGAT 1620
DB 1561 GCCTCAGATCAAAATGCAGCGGTTCTTTTGCCTAAAGGTTATTTACATTTGGCGAGTTAGAT 1620
QY 1621 TCAGATGATTATCTTGAGCTGTAGTGAACGTGTTTAAAGAAATTTTAAAGAGAT 1680
DB 1621 TCAGATGATTATCTTGAGCTGTAGTGAACGTGTTTAAAGAAATTTTAAAGAGAT 1680
QY 1681 AAAACGCTAGCTTGTGTTTATACCACTTAATAGAAACGTCAATCCGATGGTGTAAATC 1740
DB 1681 AAAACGCTAGCTTGTGTTTATACCACTTAATAGAAACGTCAATCCGATGGTGTAAATC 1740
QY 1741 GCTAATGGTTACATTTGGCGGATTTTACAGAGAAAACCTCAACGCGCTATGATGCT 1800
DB 1741 GCTAATGGTTACATTTGGCGGATTTTACAGAGAAAACCTCAACGCGCTATGATGCT 1800
QY 1801 CACCATTTTGAATTTTACGATTAGAGCTTGGCATTTAAACGGATGGATTTAAACGAAAAT 1860
DB 1801 CACCATTTTGAATTTTACGATTAGAGCTTGGCATTTAAACGGATGGATTTAAACGAAAAT 1860
QY 1861 ATTGAAAACGCGGTGATTTAGCATGTTCCCTTAACTCAGTGAAGTTGGAAAATTTTAAA 1920
DB 1861 ATTGAAAACGCGGTGATTTAGCATGTTCCCTTAACTCAGTGAAGTTGGAAAATTTTAAA 1920
QY 1921 CATCTTAAATAATCTGCTATACCGGTTATACATGGTGATACACATCCATTAAGAAA 1980
DB 1921 CATCTTAAATAATCTGCTATACCGGTTATACATGGTGATACACATCCATTAAGAAA 1980
QY 1981 CTCGGCATTTCAAAAAGAAAACCATTTTGTGTAGTCAATCAGTCAATTAATAGAACGGC 2040
DB 1981 CTCGGCATTTCAAAAAGAAAACCATTTTGTGTAGTCAATCAGTCAATTAATAGAACGGC 2040
QY 2041 ATCAATTTATTAATATGACAAAATTTGATGATTTAGATGAAGTAGAAGTATATCTTC 2100
DB 2041 ATCAATTTATTAATATGACAAAATTTGATGATTTAGATGAAGTAGAAGTATATCTTC 2100
QY 2101 AATAAACCGCTGAATATCAAGAAGAAATGGATATGTTTAAAGATCTTAAACTCATTTCAA 2160
DB 2101 AATAAACCGCTGAATATCAAGAAGAAATGGATATTTTAAAGATCTTAAACTCATTTCAA 2160
QY 2161 AATAAAGATGCAAAATGCAGTCAATTTTCTATCCCAATACATTAACGCGCTTAGTG 2220
DB 2161 AATAAAGATGCAAAATGCAGTCAATTTTCTATCCCAATACATTAACGCGCTTAGTG 2220
QY 2221 AAAAACTAAACAATATTTTGAATATAATAAAAAATATATTCGTTATTTTCTACATGTT 2280

DB 2221 AAAAACTAAACAATATTTTGAATATAATAAAAAATATATTCGTTATTTTCTACATGTT 2280
QY 2281 GATAAGAATCATCTTTACACGAGACATCAAAAAAGAAATATTTGGCTTTCTATCATAGCAC 2340
DB 2281 GATAAGAATCATCTTTACACGAGACATCAAAAAAGAAATATTTGGCTTTCTATCATAGCAC 2340
QY 2341 CAAGTGAATATTTTACTTAAATATGACATCTCATATTTACACGAGTAAATAGACATAAATAA 2400
DB 2341 CAAGTGAATATTTTACTTAAATATGACATCTCATATTTACACGAGTAAATAGACATAAATAA 2400
QY 2401 ACTGAGGCACATTTTAGTAAATATTAATAATTAAGTCAGTTAAATCTAAATTTGTAATAC 2460
DB 2401 ACTGAGGCACATTTTAGTAAATATTAATAATTAAGTCAGTTAAATCTAAATTTGTAATAC 2460
QY 2461 ATCATTTTTCATATCATGACAGCTATTCGTTAAAAATGACAGCTATGCTTTATATGAAA 2520
DB 2461 ATCATTTTTCATATCATGACAGCTATTCGTTAAAAATGACAGCTATGCTTTATATGAAA 2520
QY 2521 AAATATGATCTCGCATGAATTTTCTCAGCATTTAAACATGATTTGATCGAGAAAATCAAT 2580
DB 2521 AAATATGATCTCGCATGAATTTTCTCAGCATTTAAACATGATTTGATCGAGAAAATCAAT 2580
QY 2581 GCGCATCCACATTTTAAAAAGCTGAATTTAAACCTATTTTAAATGACATGACTTAAAGAGT 2640
DB 2581 GCGCATCCACATTTTAAAAAGCTGAATTTAAACCTATTTTAAATGACATGACTTAAAGAGT 2640
QY 2641 ATGAATGTGAAAGGGGCATCACAGGTATGTTTATGAAGTATGCGCTACCGCATGAGCTT 2700
DB 2641 ATGAATGTGAAAGGGGCATCACAGGTATGTTTATGAAGTATGCGCTACCGCATGAGCTT 2700
QY 2701 CTGACGATTTTAAAGAAAGTCATCACATCTCGGCAATCAATGATGATGTCGCCAGATAT 2760
DB 2701 CTGACGATTTTAAAGAAAGTCATCACATCTCGGCAATCAATGATGATGTCGCCAGATAT 2760
QY 2761 AACACTGAGGATTTTGGTCCCAATTTGACCTTTTAACTTTAGAAAAGAAAACCGGCCAT 2820
DB 2761 AACACTGAGGATTTTGGTCCCAATTTGACCTTTTAACTTTAGAAAAGAAAACCGGCCAT 2820
QY 2821 GTATTTTAAACATCGACCTGACTTATATGCTTTGGGAACGAAAATTTCAATTTGAGCA 2880
DB 2821 GTATTTTAAACATCGACCTGACTTATATGCTTTGGGAACGAAAATTTCAATTTGAGCA 2880
QY 2881 AATCAACAATTTCAAGTGCAAAAGGCGGAAAATATCCCGTTTAAACAAGTTCAATTAT 2940
DB 2881 AATCAACAATTTCAAGTGCAAAAGGCGGAAAATATCCCGTTTAAACAAGTTCAATTAT 2940
QY 2941 AATAGTATAACGCTATATAACATTTTGCAATTTTATTTAAAA 2979
DB 2941 AATAGTATAACGCTATATAACATTTTGCAATTTTATTTAAAA 2979

RESULT 4
US-10-011-768B-8
; Sequence 8, Application US/10011768B
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H.
; APPLICANT: DeAngelis, Paul
; APPLICANT: Kumari, Kshama
; TITLE OF INVENTION: Hyaluronan Synthase Gene and Uses Thereof
; FILE REFERENCE: 3554.011
; CURRENT APPLICATION NUMBER: US/10/011,768B
; CURRENT FILING DATE: 2001-12-11
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: US 60/064,435
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 2937
; TYPE: DNA
; ORGANISM: Pastuerella Multocida

US-10-011-768B-8

Query Match		76.7%	Score 2285;	DB 38;	Length 2937;
Best Local Similarity		86.5%	Pred. No. 0;		
Matches 2541; Conservative		0;	Mismatches 375;	Indels 21;	Gaps 1;
Qy	43	ATTTTAAAGGAAGAAATGAATACATATTATTTGAGAGTCTGCTGAAACCTAGCGGCGGAAATC	102		
Db	1	ATTTTAAAGGACAGAAATGAATACATATTATCAACAGCAATAAAGCATATAACAGCAAT	60		
Qy	103	GACTATGAATTAGCAGCTCAATATTTGAGAGTCTGCTGAAACCTAGCGGCGGAAATC	162		
Db	61	GACTATCAATTAGCAGCTCAATATTTGAGAGTCTGCGGAAATCTATGGACGGAATTT	120		
Qy	163	GTGGAATTCGAATTTATCAATGTAAAGAAAACCTCTCGACCAATTC-----	209		
Db	121	GTGGAATTCGAATTTATCAATGTAAAGAAAACCTCTCGACCAATTC-----	209		
Qy	210	-----TTATGTAAGTGAAGTAAAGAAAACAGTGTTCGCGATAGCTCATATGATATC	180		
Db	181	GCACATCTTCTGTAATAAAGAAAGAAAGTCAATGTTGCGATAGTCCGTTAGATATT	240		
Qy	262	GCAACAGCTCTTACTTTCACAGTAAAGAAAATTAACCTATCCGGAATCAGAAAAC	321		
Db	241	GCAACAGCTCTTACTTTCACAGTAAAGAAAATTAACCTATCCGGAATCAGAAAAC	300		
Qy	322	AGTTTAAAGAAATGAATGAAATCTATCACCTGGGAAAATCGGAGAACGCAAGAAATCAGA	381		
Db	301	AGTTTAAAGAAATGAATGAAATCTATCACCTGGGAAAATCGGAGAACGCAAGAAATCAGA	360		
Qy	382	ANGGTGAAGTACTGACCAAGATTTTCCCTAAAGATTTGTTCTGCTGCCATTTCCAGAT	441		
Db	361	CGGTCGCCCTTGTACCAAGAAATTTTCCCAAGATCTGTTTGTAGCGCCTTTACCTGAT	420		
Qy	442	CATGTTAATGATTTACATGGTACAAAATCGGAAAAGAGCTAGGTATPAAAGCCGTGA	501		
Db	421	CATGTTAATGATTTACATGGTACAAAAGCGAAAAGAGACTTGGCATAAAACCTGAA	480		
Qy	502	AATGAAGATATCGGCTCTTCTATTATTATTCCTACATTTAATCTGAGCGGTATTTAGAT	561		
Db	481	CATCAACATGTTGGCTCTTCTATTATTATTCCTACATTTAATCTGAGCGGTATTTAGAT	540		
Qy	562	ATACGTTAGCTGTTTGGTCAATCAGAAAACAACTACCCATTTGAGTGTGTTGCTGCA	621		
Db	541	ATACATTTAGCTGTTTGGTCAATCAGAAAACAACTACCCATTTGAGTGTGTTGCTGCA	600		
Qy	622	GATGATGGTAAAGAAATCTACTTACCAATGTCGAAAATACGAAACAAAACCTTGAC	681		
Db	601	GATGATGGTAAAGAAATCTACTTACCAATGTCGAAAATACGAAACAAAACCTTGAT	660		
Qy	682	ATAAGATATGTAACACAAAAGATTTATGGATATCAATTTGTCGAGTCAGAAACTTAGGT	741		
Db	661	ATTCGCTACGTCAGACAAAAGATTAACGGTTTTCACGCGAGTCCGCTCGGAATATGGGA	720		
Qy	742	TTACGTACAGCAAAATGATTTTGTCTCGATTTCTAGCATTCGATATGGCACCAACAA	801		
Db	721	TTACGCTTACGCAAAATGATTTTATTTAGTCTTCTAGCATTCGATATGGCACCAATCCA	780		
Qy	802	TTATGGGTTCAATCTTATCTACAGAACTATTAGAAGCAATGATATGTTTAAATGGA	861		
Db	781	TTATGGGTTCAATCTTATGTTGAGAGCTATTAGAAGATGATGATTTAAACATCATTTGT	840		
Qy	862	CCTAGAAAATATGGGATCTACTAATATTACCGCAGAACTTCCCTTAAACGATCCATAT	921		
Db	841	CCAAGAAAATACATCGATACACACATATTGACCCCAAGAACTCTTAAATTAACCGAGT	900		
Qy	922	TTAATAGAATCACTACCTGAAACCGCTACAAATTAACAACTCTCGATTTACATCAAAAGGA	981		
Db	901	TTGCTTGAATCATTTACCAAGAGTGAACCAATTAATAGTGTGCGGCAAAAGGGGAAGGA	960		
Qy	982	AATATATCTGTTGGATTTGGAGATTTAGACATTTTCAAAAACCGGATTAATCTACGCTATGT	1041		
Db	961	ACAGTTTCTGCTGCGCTTGAACAAATTCGAAAAAAGAGTATTTTCAATTAACCGCTGAATCA	1020		

Qy	1042	GATTCCTCGTTTCGTTATTTTGTTCGGGTAAATGTTTCATTTTCTTAAAGAAATGGCTAAAT	1101
Db	1021	GATTCGCTTTTCGTTTTCGGGTAAATGTTTCGCTTAAAGAAATGGCTAAAT	1080
Qy	1102	AAAGTAGGTTGGTTCGATGAAGAATTTAATCATTTGGGGGGGGAAGATGTAAGATTTGGT	1161
Db	1081	AAATCCGGTTTCTTTGATGAGGAATTTAATCACTGGGTGGAGAAGATGTAAGATTTGGA	1140
Qy	1162	TACAGATTTATTTGCCAAAGGCTGTTTTTCAGAGTAATTTGACGGCGGAATGGCCATCCAT	1221
Db	1141	TATCGCTTTATTTCCGTTACGGTAGTTTCTTTAAACATATTGATGGCATTTATGCCCTACCAT	1200
Qy	1222	CAAGAACCCCTTGGTAAAGAAAAATCAACAGAACGCGAAGCTGGTAAAAAGTATTACGGCTT	1281
Db	1201	CAAGACCCACCGGTTAAGAAAAATCAACCGGATCGTGAAGCGGGAAGAAATATTACGGCTC	1260
Qy	1282	AAAATTTGGAAGAAAAGGTACCTTTACATCTATAGAAGCTTTTACCANAGAAGATTCGA	1341
Db	1261	GATATTATGAGAAAAAGGTCCCTTATATCTATAGAAAACTTTTACCANAGAAGATTCG	1320
Qy	1342	CATATTTCATAGAAATACCTTTAGTTTCTATTTATATCCCGCTTATAACTGTGCAATTTAT	1401
Db	1321	CATATTTCATAGAAATACCTTTAGTTTCTATTTATATCCCGCTTATAACTGTGCAATTTAT	1380
Qy	1402	ATTTCAAGATGTTAGATAGTCTCTTAATCAAACTGTTTGTCTGATCTCCGAGGTTTGTATT	1461
Db	1381	ATTTCAAGGTTGCGTAGATAGTCTGATCAATCAAGCTGTTTGTGATCTCCGAGGTTTGTATT	1440
Qy	1462	TGTAACGATGTTCAACAGATAAATACCTTTAGAAAGTATCAATAAGCTTTTATGGTAAATAT	1521
Db	1441	TGTAACGATGTTCAACAGATAAATACCTTTAGAAAGTATCAATAAGCTTTTATGGTAAATAT	1500
Qy	1522	CCTAGGGTACGATCATGCTTAAACCAATGCGGAATAGCCTCAGCATCAATATGCGGCC	1581
Db	1501	CCTAGGGTACGATCATGCTTAAACCAATGCGGAATAGCCTCAGCATCAATATGCGGCC	1560
Qy	1582	GTTCTTTTGTGCTTAAAGGTTATTACATTTGGCAGTTTAGATTTCAGATGATTATCTTGAGCCT	1641
Db	1561	GTTCTTTTGTGCTTAAAGGTTATTACATTTGGCAGTTTAGATTTCAGATGATTATCTTGAGCCT	1620
Qy	1642	GATCAGTTGAACTGTGTTTAAAGAAATTTTAAAGATATAAACGCTAGCTTGTGTTTAT	1701
Db	1621	GATCAGTTGAACTGTGTTTAAAGAAATTTTAAAGATATAAACGCTAGCTTGTGTTTAT	1680
Qy	1702	ACCCTAATAGAAAGCTCAATCCGGATGCTAGCTTTAATCGCTTAATGGTTTACAAATGGCCA	1761
Db	1681	ACCCTAATAGAAAGCTCAATCCGGATGCTAGCTTTAATCGCTTAATGGTTTACAAATGGCCA	1740
Qy	1762	GAATTTTCCAGGAAAAACTCAACACGGCTATGATTGCTCAGCATTTTGAATGTTTAGG	1821
Db	1741	GAATTTTCCAGGAAAAACTCAACACGGCTATGATTGCTCAGCATTTTGAATGTTTAGG	1800
Qy	1822	ATTAGAGCTTGGCATTTTAAACGGATGATTTAAACGAAAAATATTGAAACCGCGTGGATAT	1881
Db	1801	ATTAGAGCTTGGCATTTTAAACGGATGATTTAAACGAAAAATATTGAAACCGCGTGGATAT	1860
Qy	1882	GACATGTTCTTAAACTCAGTCAAGTGGAAAAATTTAAACATCTTAAATAAAATCTGCTAT	1941
Db	1861	GACATGTTCTTAAACTCAGTCAAGTGGAAAAATTTAAACATCTTAAATAAAATCTGCTAT	1920
Qy	1942	AACCCGCTATTACATGGTGAATACACATCCATTAAGAAAACTGGGCATTTCAAAAGAAAAAC	2001
Db	1921	AACCCGCTATTACATGGTGAATACACATCCATTAAGAAAACTGGGCATTTCAAAAGAAAAAC	1980
Qy	2002	CATTTTGTGTAGTCAATCAGTCATTTAAATAGCAAGGATCAATTTATTTATTTATGAC	2061
Db	1981	CATTTTGTGTAGTCAATCAGTCATTTAAATAGCAAGGATCAATTTATTTATTTATGAC	2040
Qy	2062	AAATTTGATGATTTAGATGAAAGTAGAAGTATATCTTCAATTAACCCCTGAATATCAA	2121
Db	2041	GAATTTGATGATTTAGATGAAAGTAGAAGTATATTTTCAATTAACCCCTGAATATCAA	2100

Db 901 TTGCTTGAATCATACCAGAGTGAACCAATTAATAGTGTGGCGCAAAAGGGGAAGGA 960
Qy 982 AATATATCGTTGGATTCAGAGATTAAGACATTTCAAAACCGGATTAATCTACGCTATGT 1041
Db 961 ACAGTTTCTCGGATTCGGCGTTAGAACAAATTCGAAAACAGAAAATCTCCGCTTATCC 1020
Qy 1042 GATTCTCCGTTTCGTTATTTTGTTCGGGTAATGTTCGATTTTCTTAAAGATGGCTAAAT 1101
Db 1021 GATTCCGCTTTCGCTTTTTCGGCGGGTAATGTTCGCTTAAAGATGGCTAAAT 1080
Qy 1102 AAGTAGTGGTTTCGATGAAGAATTAATCATTTGGGGGGCGGAGAGTGAATTTGGT 1161
Db 1081 AAATCCGGTTTCTTTGTATGAGAAATTAATCACTGGGTGGAGAGATGTGGAATTTGGA 1140
Qy 1162 TACAGATTATTTGCAAGGCTGTTTTTTCAGAGTAATTTGACGGGGAATGGCCATCCAT 1221
Db 1141 TATCGCTTATTCGCTTACGGTAGTTTCTTTTAAACATATTGATGGCATTTATGGCTTACCAT 1200
Qy 1222 CAAGAACCCACTGGTAAAGAAAATGAACAGACCGGAAGCTGGTAAAGTATTTACGGTT 1281
Db 1201 CAAGAGCCACCAGGTAAAGAAAATGAACCCGATCGTGAAGCGGGAAGAAAATATTACGCTC 1260
Qy 1282 AAAATTTGAAAGAAAGGTACCTTACATCTATGAAGAGCTTTTACCAATAGAAGATTCA 1341
Db 1261 GATATTATGAGAGAAAAGGTCCCTTATATCTATAGAAAACCTTTTACCAATAGAAGATTTCG 1320
Qy 1342 CATATTTCATAGAAATACCTTTAGTTTCTATTTATATCCCGCTTATTAAGTGTGCAAAATAT 1401
Db 1321 CATATCAATAGAGTACCTTTAGTTTCTATTTATATPCCAGCTTATTAAGTGTGCAAAATAT 1380
Qy 1402 ATTCAAAAGATGTAGATAGTGCCTTTAATCAAACTGTGTGCGATCTCGAGGTTTGTATT 1461
Db 1381 ATTCACAGTTTCGTTAGATAGTGCCTGAATCAGACTGTGTGTGATCTCGAGGTTGTATT 1440
Qy 1462 TGTAAAGATGTTCAACAGATAATACCTTAGAAGTGAATCAATAGCTTTATGGTAATAT 1521
Db 1441 TGTAAAGATGTTCAACAGATAATACCTTAGAAGTGAATCAATAGCTTTATGGTAATAT 1500
Qy 1522 CCTAGGTGACCATCATCTCTAAACCAATGCGGAATAGCTTCAGCATCAATGCGGCC 1581
Db 1501 CCTAGGTGACCATCATCTCTAAACCAATGCGGAATAGCTTCAGCATCAATGCGGCC 1560
Qy 1582 GTTCTTTTGTAAAGGTTATTACATTTGGCGAGTTAGATTTCAGATGATTATCTTGAGGCT 1641
Db 1561 GTTCTTTTGTAAAGGTTATTACATTTGGCGAGTTAGATTTCAGATGATTATCTTGAGGCT 1620
Qy 1642 GATCGAGTTGAACGTGTTTAAAGAAATTTTAAAGATAAAGCCTAGCTGTGTTTAT 1701
Db 1621 GATCGAGTTGAACGTGTTTAAAGAAATTTTAAAGATAAAGCCTAGCTGTGTTTAT 1680
Qy 1702 ACCACTAATAGAAACGTTCAATCCGATGGTAGCTTAATCGCTTAATGTTTACAAATGGCCA 1761
Db 1681 ACCACTAATAGAAACGTTCAATCCGATGGTAGCTTAATCGCTTAATGTTTACAAATGGCCA 1740
Qy 1762 GAATTTTTCACGAGAAAACCTCACACGCTATGATGCTCACCATTTTAGAGATGTTTACG 1821
Db 1741 GAATTTTTCACGAGAAAACCTCACACGCTATGATGCTCACCATTTTAGAGATGTTTACG 1800
Qy 1822 ATTAGAGCTTGGCATTTAACCGATGGATTTAACGAAATATTGAACCGCGCTGATTTAT 1881
Db 1801 ATTAGAGCTTGGCATTTAACCTGATGGATTTCAATGAAAATTTGAAAATGCGGTAGACTAT 1860
Qy 1882 GACATGTTCTTAACTCAGTGAAGTTGGAAAATTTAAACATCTTAATAAATCTGCTAT 1941
Db 1861 GACATGTTCTTAACTCAGTGAAGTTGGAAAATTTAAACATCTTAATAAATCTGCTAT 1920
Qy 1942 AACCGCGTATTACATGTTGATACACATCTTAAAGAACTCGGATTTCAAAAGAAAAC 2001
Db 1921 AACCGGTGATTACATGTTGATACACATCTTAAAGAACTCGGATTTCAAAAGAAAAC 1980
Qy 2002 CATTTTGTGTAGTCAATCAGTCAATTAATAGAACGCAATTAATTAATTAATGAC 2061
Db 1981 CATTTTGTGTAGTCAATCAGTCAATTAATAGAACGCAATTAATTAATTAATGAC 2040

Qy 2062 AAATTTGATGATTAGATGAAAGTATATCTTCAATAAAACCCGCTGAATATCAA 2121
Db 2041 GAATTTGATGATTAGATGAAAGTATATCTTCAATAAAACCCGCTGAATATCAA 2100
Qy 2122 GAAGAAATGGATATGTTTAAAGATCTTAAACATCTTCAAAATAAAGATGCCAAAATCGCA 2181
Db 2101 GAAGAGATTGATATCTTAAAGATATTAATAATCATCCAGATAAAGATGCCAAAATCGCA 2160
Qy 2182 GTCAGTATTTCTATCCCAATACATTAACGGCTTAGTGAAGAACTAAACATATTTAT 2241
Db 2161 GTCAGTATTTCTATCCCAATACATTAACGGCTTAGTGAAGAACTAAACATATTTAT 2220
Qy 2242 GAATATAATAAAATATATTCGTTTATTTTACATGTTTGAAGAACTATCTTTACACCA 2301
Db 2221 GAATATAATAAAATATATTCGTTTATTTTACATGTTTGAAGAACTATCTTTACACCA 2280
Qy 2302 GACATCAAAAAGAAATATTTGGCTTTCTATCATCAAGCACCAAGTGAATATTTTACTAAAT 2361
Db 2281 GATATCAAAAAGAAATACTAGCTTTCTATCATCAACATCAAGTGAATATTTTACTAAAT 2340
Qy 2362 AATGACATCTCATTTTACAGAGTATAGACTAATAAAGCTGAGGACATTTTAAAGTAAT 2421
Db 2341 AATGATATCTCATTTTACAGAGTATAGACTAATAAAGCTGAGGACATTTTAAAGTAAT 2400
Qy 2422 ATTAATAAATTAAGTCAGTTAAATCTTAAATTTGTAATACATCATTTTGTATATCATGAC 2481
Db 2401 ATTAATAAATTAAGTCAGTTAAATCTTAAATTTGTAATACATCATTTTGTATATCATGAC 2460
Qy 2482 AGCCTATTCGTTTAAAGTACAGCTATGCTTATATGAAAAAATATATGATGTCGCGCATGAAT 2541
Db 2461 AGCCTATTCGTTTAAAGTACAGCTATGCTTATATGAAAAAATATATGATGTCGCGCATGAAT 2520
Qy 2542 TTCTCAGCATTAACACATGATTTGGATCGAGAAAAATCAATGCGCATCCACCATTTAAAG 2601
Db 2521 TTCTCAGCATTAACACATGATTTGGATCGAGAAAAATCAATGCGCATCCACCATTTAAAG 2580
Qy 2602 CTGATTAACACCTATTTTAAAGCAATGACTTAAAGAAATGATGAATCTGAAAGGGCATCA 2661
Db 2581 CTCATTAACACCTATTTTAAAGCAATGACTTAAAGAAATGATGAATCTGAAAGGGCATCA 2640
Qy 2662 CAAGTATGTTTATGAAGTATGCGTACGCGCATGAGCTTCTGACGATTTTAAAGAGTC 2721
Db 2641 CAAGTATGTTTATGAAGTATGCGTACGCGCATGAGCTTCTGACGATTTTAAAGAGTC 2700
Qy 2722 ATCAATCTCGCAATCAATTTGATAGTGTGCGAGAAATATAACACTGAGGATTTGGTTC 2781
Db 2701 ATCAATCTCGCAATCAATTTGATAGTGTGCGAGAAATATAACACTGAGGATTTGGTTC 2760
Qy 2782 CAATTTGCACTTTTAACTTTAGAAAAGAAACCGGCGCATGTTTAAATAAACATCGACC 2841
Db 2761 CAATTTGCACTTTTAACTTTAGAAAAGAAACCGGCGCATGTTTAAATAAACATCGACC 2820
Qy 2842 CTGACTTATATGCTTTGGGAGCAAAATTTACAATGGCAAAATGAACAAATTTCAAAGTGCA 2901
Db 2821 CTGACTTATATGCTTTGGGAGCAAAATTTACAATGGCAAAATGAACAAATTTGAAAGTGCA 2880
Qy 2902 AAAAAAGGCAAAATATCCCGTTAAACAAGTTCAATTTAATAAGTATAACGCTATAA 2958
Db 2881 AAAAAAGGCAAAATATACCTGTTAAACAAGTTCAATTTAATAAGTATAACGCTATAA 2937

RESULT 6

PCT-US99-07289-2

; Sequence 2, Application PC/TUS9907289

; GENERAL INFORMATION:

; APPLICANT: Board of Regents of the University of Oklahoma

; TITLE OF INVENTION: NUCLEIC ACID ENCODING HYALURONAN SYNTHASE AND METHODS

; TITLE OF INVENTION: OF USE

; FILE REFERENCE: 617022-7

; CURRENT APPLICATION NUMBER: PCT/US99/07289

; CURRENT FILING DATE: 1999-04-01

; EARLIER APPLICATION NUMBER: 60/080,414

; EARLIER FILING DATE: 1998-04-02									
; EARLIER APPLICATION NUMBER: 60/178,851									
; EARLIER FILING DATE: 1998-10-26									
; NUMBER OF SEQ ID NOS: 29									
; SOFTWARE: PatentIn Ver. 2.0									
; SEQ ID NO 2									
; LENGTH: 2937									
; TYPE: DNA									
; ORGANISM: Pasteurella multocida									
PCT-US99-07289-2									
Query Match									
Best Local Similarity 76.6%; Score 2283.4; DB 1; Length 2937;									
Matches 2540; Conservative 0; Mismatches 376; Indels 21; Gaps 1;									
Qy	43	ATTTTAAAGGAAGAAATGAATACATATATACACAGCAATAAAGCATATACAGCAAT	102						
Db	1	ATTTTAAAGGACAGAAATGAATACATATATACAGCAATAAAGCATATACAGCAAT	60						
Qy	103	GACTATGAATTAGCACTCAAAATATTGTGAGAGTCTGCTGAACCTACGGCGGAAATC	162						
Db	61	GACTATCAATTAGCACTCAAAATATTGTGAAAGTCGGCGGAATCTATGACGGAAAT	120						
Qy	163	GTGGAATCCAAATATCAATGTAAAGAAAACTCTGACCAATTC-----	209						
Db	121	GTGGAATTCAAATACCAATGCCAAGAAACTCTCAGCACATCCTTCTGTTAATCA	180						
Qy	210	-----TTATGTAAGTGAAGATAAATAAAGAGTGTGGATAGCTCATTAGATATC	261						
Db	181	GCACATCTTCTCTAAATAAAGAAAGAAAGTCAATGTTGGATAGTCGGTATAGAT	240						
Qy	262	GCACACAGCTCTTACTTCCACGTAATAAATAAATTAACCTATCCGAATCAGAAAAAC	321						
Db	241	GCACACAACTGTTACTTCCACGTAATAAATAAATTAAGTACTTCTGACCTCGGAAAC	300						
Qy	322	AGTTTAAAAATAAATGGAATCTATCACTCGGGAATAATCGAGAACCGCAGAAATCAGA	381						
Db	301	ACGTTAAAAATAAATGGAATGCTCACTGAGAATAATCTGAAATCGGGAGGTAAGA	360						
Qy	382	AGGTGGAAGTACTACCAAGATTTTCTTAAGATCTGTTCTTGTCTCCATTCGCCAGAT	441						
Db	361	CGGTCGCCCTTTACCAAAAGATTTTCCCAAGATCTGGTTTACGCCCTTTACCTGAT	420						
Qy	442	CATGTTAATGATTTTACATGGTCAAAAATCGAAAAAAGCTTAGGTATAAAGCCTGTA	501						
Db	421	CATGTTAATGATTTTACATGGTACAAAAGCGGAAGAAGACTTGGCATATAAACCTGAA	480						
Qy	502	AATAAGAATATCGGCTCTTCTATATTATTCCTACATTTAATCGTAGCCGTATTTAGAT	561						
Db	481	CATCAACATGTTGGTCTTCTATATGTTTACAACTTCAATCGACGCAATTTTATCG	540						
Qy	562	ATAACGTTAGCCTGTTGGTCAATCAAGAAACAACTTACCATTGGAAGTCGTTGTGCA	621						
Db	541	ATTACATATGCCCTTTAGTAACCAAAAACACATTTACCCTGTTGAAGTTATCGTGACA	600						
Qy	622	GATGATGGTAGTAAGAAAACTTACTTACCATTTGTCAAAAATAACGAACAAAACTTGAC	681						
Db	601	GATGATGGTAGTCAGGAAGATCATCACCGATCATTCGCCAATATGAAATAAATTTGGAT	660						
Qy	682	ATAAGTATGTAGACAAAAAGATTTATGGATATCAATTTGTGCGAGTCAGAACTTAGT	741						
Db	661	ATTGCGTACGTCAGACAAAAAGATTAACGGTTTCAAGCCAGTCGCCGTGCGAATATGGGA	720						
Qy	742	TTAGTCACAGCAAGATGATTTTGTCTCGATCTAGACTGCGATATGGCACCAACAACA	801						
Db	721	TTAGCCTTAGCAAAATATGACTTTATTTGGCTTACTCGACTGTGATATGGCCCAATCCA	780						
Qy	802	TTATGGGTTTCATTTCTTCTACAGAACTATTTAGAGCAAGATGATTTGTTTAAATGGA	861						
Db	781	TTATGGGTTTCATTTCTTGTGACAGAGCTATTTAGAAGATGATGATTTTAACAATCAT	840						
Qy	862	CCTAGAAAAATATGTTGATCTCAATATTTACCGCAGCAATTTCCCTTAACGATCCATAT	921						

Db 841 CCAAGAAAATACATCGATACACAAACATATTGACCCAAAAGACTTCTTAATAAACGGAGT 900
QY 922 TTAATAGATCACTACCTGAACCCGCTACAATAACAATCCCTCGATTACATCAARAAGA 981
Db 901 TTGCTTGAATCATATTACAGAAGTGAAGAACCAATAATAGTTGGCCGCAAAAGGGGAAGA 960
QY 982 AATATATCGTTGGATTGGAGATTAGAACATTTTCAAAAAACCCGATATATCTACGCTCTATGT 1041
Db 961 ACAGTTCTCTGGATTGGCGCTTAGAACAAATTCGAAAAACAGAAAATCTCCGCTTATCC 1020
QY 1042 GATTTCCGTTTCGTTATTTTGTGGGGTAATGTTGCAATTTTCTAAGAATGGCTAAAT 1101
Db 1021 GATTCGCTTTCGTTTTCGCGGGGTAATGTTGCTTTCGCTAAAAAATGGCTAAAT 1080
QY 1102 AAGTAGTTGGTTTCGATGAAGAATTTAATCATTTGGGGGGCGAAGATGTFAGAATTTGGT 1161
Db 1081 AAATCGGGTTCTTTGATGAGGAATTTAATCACTGGGGTGGAGAAGATGFGGAATTTGGA 1140
QY 1162 TACAGATTATTTGCCAAAGGCTGTTTTTTCAGAGTAATTTGACGCGCGGAATGGCCATCCAT 1221
Db 1141 TATCGCTTATTCGTTTACGGTAGTTCTTTAAACTATTGATGGCATTTATGGCTTACCAT 1200
QY 1222 CAAGAACCACCTGGTAAAGAAAATGAACAGACGGAGCTGGTAAAGTATTTACGCTT 1281
Db 1201 CAAGAGCCACCGTAAAGAAAATGAACCGGATCGTGAAGCGGGAATAATATTACGCTC 1260
QY 1282 AAAATTGTGAAGAAAGGTACCTTACATCTATAGAAGCTTTTACCATAAGAGATTCA 1341
Db 1261 GATATTAGAGAAAGGTCCCTTATCTATAGAAAACCTTTTACCATAAGAGATTGCG 1320
QY 1342 CATATTCATAGAAATACCTTTAGTTCTTATTTATATCCCGCTTATATACTGTGCAAAATAT 1401
Db 1321 CATATCAATAGAGTACCTTTAGTTTCAATTTATATCCAGCTTATACTGTGCAAACTAT 1380
QY 1402 ATTCAAAGATGTGATAGTAGTCTTAAACAACCTGTTGCGATCTCGAGGTTTGTATT 1461
Db 1381 ATTCAGGTTGCGTAGAGTAGTGCCTGAATCAGACTGTTGTTGATCTCGAGGTTTGTATT 1440
QY 1462 TGTAAACATGGTTCAACAGATAATACCTTAGAAGTGTATCAATAAGCTTTTATGTAATAT 1521
Db 1441 TGTAAACATGGTTCAACAGATAATACCTTAGAAGTGTATCAATAAGCTTTTATGTAATAT 1500
QY 1522 CCTAGGTTAGCATCATGCTTAACCAAAATGCGCGGATAGCCTCAGCATCAAAATGCGACC 1581
Db 1501 CCTAGGTTAGCATCATGCTTAACCAAAATGCGCGGATAGCCTCAGCATCAAAATGCGACC 1560
QY 1582 GTTTCTTTTGTAAAGTTATTACATTGGGCGAGTTAGATTTCAGATGATTATCTTGAGCCT 1641
Db 1561 GTTTCTTTTGTAAAGTTATTACATTGGGCGAGTTAGATTTCAGATGATTATCTTGAGCCT 1620
QY 1642 GATCAGTTGAACGTGTGTTTAAAGAAATTTTAAAGATAAAACGCTAGCTTGTGTTTAT 1701
Db 1621 GATCAGTTGAACGTGTGTTTAAAGAAATTTTAAAGATAAAACGCTAGCTTGTGTTTAT 1680
QY 1702 ACCACTAATAGAAAGCTCAATCCGGATGGTAGCTTTAATCGCTAATGGTTACAATGGCCA 1761
Db 1681 ACCACTAATAGAAAGCTCAATCCGGATGGTAGCTTTAATCGCTAATGGTTACAATGGCCA 1740
QY 1762 GAATTTTCCAGAGAAAACTCACAAACGCTATGATTGCTCACCATTATTTAGAAATGTTTACG 1821
Db 1741 GAATTTTCCAGAGAAAACTCACAAACGCTATGATTGCTCACCATTATTTAGAAATGTTTACG 1800
QY 1822 ATTAGAGCTTGGCATTTAACGGATGGATTTAACGAAAATATTGAAAACGCGGTGGAATTAT 1881
Db 1801 ATTAGAGCTTGGCATTTTAACGTGATGATTCATGAAAATTTGAAAATGCGGTGAGCATAT 1860
QY 1882 GACATGTTCCCTTAACTCAGTGAAGTGTGAAAATTTTAAACATCTTAAATAAAATCTGCTAT 1941
Db 1861 GACATGTTCCCTTAACTCAGTGAAGTGTGAAAATTTTAAACATCTTAAATAAAATCTGCTAT 1920
QY 1942 AACCCGGTATTACATGGTGATAAACATCCATTAAGAAACCTGCGATTCAAAAGAAAAAC 2001
Db 1921 AACCCGTATTACATGGTGATAAACATCAATTAAGAAACCTTGGCATTTCAAAAGAAAAAC 1980

QY 2002 CATTTTGTGTAGTCAATCAGTCAATTAATAAGACAGGCATCAATTTATTAATATTATGAC 2061
Db 1981 CATTTTGTGTAGTCAATCAGTCAATTAATAAGACAGGCATCAATTTATTAATATTATGAC 2040
QY 2062 AAATTTGATGATTTAGATGAAGTAGAAGATATATCTTCAATAAAACCGCTGAATATCAA 2121
Db 2041 GAATTTGATGATTTAGATGAAGTAGAAGATATTTTCAATAAAACCGCTGAATATCAA 2100
QY 2122 GAAGAATGATATGTTAAAGATCTTAAAGACTTCAATCAAAATAAAGATGCCAAATCGCA 2181
Db 2101 GAAGATTTGATATCTTAAAGATATTTAAATCATCCAGATAAGATGCCAAATCGCA 2160
QY 2182 GTCAGTATTTTATCCCAATACATTAACCGGCTTAGTGA AAAACCTTAACAATATTATT 2241
Db 2161 GTCAGTATTTTATCCCAATACATTAACCGGCTTAGTGA AAAACCTTAACAATATTATT 2220
QY 2242 GAATATTAATAAATAATATATTTCGTTATTATTCTACATCTGTATAGAATCATCTTTACACCA 2301
Db 2221 GAATATTAATAAATAATATATTTCGTTATTATTCTACATCTGTATAGAATCATCTTTACACCA 2280
QY 2302 GACATCAAAAAGAAATATTGGCTTTCTATCATAGCACCAAGTGAATTTTACTAAAT 2361
Db 2281 GATATCAAAAAGAAATATTAGCTTCTATCATATAACATCAAGTGAATTTTACTAAAT 2340
QY 2362 AATGACATCTCATATTTACAGAGTAATAGACTAATAA AAACTGAGGCACATTTAAGTAAT 2421
Db 2341 AATGATATCTCATATTACACGAGTAATAGATTAAATAA AAACTGAGGCATTTAAGTAAT 2400
QY 2422 ATTAATAAATTAAGTCAAGTTAAATCTAAATGTGAATACATCATTTTGTATAAATCATGAC 2481
Db 2401 ATTAATAAATTAAGTCAAGTTAAATCTAAATGTGAATACATCATTTTGTATAAATCATGAC 2460
QY 2482 AGCCTATTTCGTTAAAAATGACAGCTATGCTTATATGAAAAAATATGATGTGGCATGAAT 2541
Db 2461 AGCCTATTTCGTTAAAAATGACAGCTATGCTTATATGAAAAAATATGATGTGGCATGAAT 2520
QY 2542 TTCACGATTTAAACACATGATTGGATCGAGAAATCAATCGCATCCACCATTTTAAAAAG 2601
Db 2521 TTCACGATTTAAACACATGATTGGATCGAGAAATCAATCGCATCCACCATTTTAAAAAG 2580
QY 2602 CTGATTAACACCTTTTAAATGACAATGACTTAAAGAACTATGAATGTGAAAGGGGCATCA 2661
Db 2581 CTCATTAACACCTTTTAAATGACAATGACTTAAAGAACTATGAATGTGAAAGGGGCATCA 2640
QY 2662 CAAGTATGTTTATGAAGTATGCGCTACCGCATGAGCTTCTGACGATTTTAAAGAAGTC 2721
Db 2641 CAAGTATGTTTATGAAGTATGCGCTACCGCATGAGCTTCTGACGATTTTAAAGAAGTC 2700
QY 2722 ATCATCTCTGCCAATCAATTTGATGTGCGAGAAATATAACACCTGAGGATATTTGGTTC 2781
Db 2701 ATCATCTCTGCCAATCAATTTGATGTGCGAGAAATATAACACCTGAGGATATTTGGTTC 2760
QY 2782 CAATTTGCACCTTTTAACTTTAGAAAAAGAAACCGGCCATGTTATTTAATAAACAATCGACC 2841
Db 2761 CAATTTGCACCTTTTAACTTTAGAAAAAGAAACCGGCCATGTTATTTAATAAACAATCGACC 2820
QY 2842 CTGACTTATATGCTTTGGAGCAAAAATTTACAATGACAAAATGAACAAATTTCAAAGTGA 2901
Db 2821 CTGACTTATATGCTTTGGAGCAAAAATTTACAATGACAAAATGAACAAATTTCAAAGTGA 2880
QY 2902 AAAAAAGCGAAAATATATCCCGTTAAACAGTTTCAATTTAATAGTATAACGCTATAA 2958
Db 2881 AAAAGGAGAAAATATATACCTCTTAACAAGTTTCAATTTAATAGTATAACTCTATAA 2937

RESULT 8

US-09-469-200-2

; Sequence 2, Application US/09469200

; GENERAL INFORMATION:

; APPLICANT: Board of Regents of the University of Oklahoma

; TITLE OF INVENTION: NUCLEIC ACID ENCODING HYALURONAN SYNTHASE AND METHODS

; TITLE OF INVENTION: OF USE

1942 AACCCGGTATTACATGGTGATTAACACATCCATTAAGAACTCGGCATTCATAAAGAAAC 2001
1921 AACCCGTATTACATGGTGATTAACACATCAATTAAGAACTTTGGCATTCATAAAGAAAC 1980
2002 CATTGTGTAGTCAATCAGTCATTAAATACACAGGCTCAATTAATTAATTAATATGAC 2061
1981 CATTGTGTAGTCAATCAGTCATTAAATACACAGGCTCAATTAATTAATTAATATGAC 2040
2062 AAATTTGATGATTAGATGAAAGTAGAAGTATATCTTCAATTAACCCGCTGAATATCAA 2121
2041 GAATTTGATGATTAGATGAAAGTAGAAGTATATCTTCAATTAACCCGCTGAATATCAA 2100
2122 GAAGAAATGATATGTTAAAGATCTTAACCTCAATCAAAATAAGATCCCAAAATCGCA 2181
2101 GAAGAGATTGATATCTTAAAGATATTAAGATCATCCAGATAAGATGCCAAATCGCA 2160
2182 GTCAATTTTCTATCCCAATACATTAACCGCTTAGTGAAAGAACTAAACAAATATATT 2241
2161 GTCAATTTTCTATCCCAATACATTAACCGCTTAGTGAAAGAACTAAACAAATATATT 2220
2242 GAATATAATAAATAATATCGTTTCTATACATGTTGATAAGATCAATCATCTTACACCA 2301
2221 GAATATAATAAATAATATCGTTTCTATACATGTTGATAAGATCAATCATCTTACACCA 2280
2302 GACATCAAAAGAAATATTGGCTTCTATACATGACCAAGTGAATATTTTACTAAAT 2361
2281 GATATCAAAAGAAATATAGCTTCTATACATGACCAAGTGAATATTTTACTAAAT 2340
2362 AATGACATCTCATATTACAGGATATAGACTAATAAAGCTGAGGACATTTTAAAT 2421
2341 AATGATATCTCATATTACAGGATATAGATTAAATAAAGCTGAGGACATTTTAAAT 2400
2422 ATTAATAAATAAGTCAGTTAAATCTAAATTTGAATACATCTTTTGAATATCATGAC 2481
2401 ATTAATAAATAAGTCAGTTAAATCTAAATTTGAATACATCTTTTGAATATCATGAC 2460
2482 AGCTTATCGTTTAAATGACAGTATGCTTATATAAAGAAATATGATGCGGATGAAT 2541
2461 AGCTTATCGTTTAAATGACAGTATGCTTATATAAAGAAATATGATGCGGATGAAT 2520
2542 TTCTCAGCATTAACATGATTTGATGAGGAGAAATCAATCGGATCCACCATTTAAAG 2601
2521 TTCTCAGCATTAACATGATTTGATGAGGAGAAATCAATCGGATCCACCATTTAAAG 2580
2602 CTGATTAACACCTATTTTATGACATGACTTAAGAGTATGATGAAAGGGCATCA 2661
2581 CTGATTAACACCTATTTTATGACATGACTTAAGAGTATGATGAAAGGGCATCA 2640
2662 CAAGGTATGTTTATGAGTATGCGCTACCGCATGAGCTTCTGACGATTTAAAGAGTC 2721
2641 CAAGGTATGTTTATGAGTATGCGCTACCGCATGAGCTTCTGACGATTTAAAGAGTC 2700
2722 ATCAATCTCGCCAAATCAATGATGATGTCGAGAAATATACACTGAGGATTTTGGTTC 2781
2701 ATCAATCTCGCCAAATCAATGATGATGTCGAGAAATATACACTGAGGATTTTGGTTC 2760
2782 CAATTTGCACTTTTAACTTTAGAAAGAAACCGCCATGATTTTAAATAAATCAATCGAC 2841
2761 CAATTTGCACTTTTAACTTTAGAAAGAAACCGCCATGATTTTAAATAAATCAATCGAC 2820
2842 CTGACTTTATGCTTGGGACGAAATTAACAATGGAACAAATGAACAAATTTCAAAGTGA 2901
2821 CTGACTTTATGCTTGGGACGAAATTTACAATGGAACAAATGAACAAATTTCAAAGTGA 2880
2902 AAAAAGGCGAAATATCCCGTTTAAACAGTTCAATTTATATATATTAACCTATAA 2958
2881 AAAAAGGCGAAATATACCTGTTTAAACAGTTCAATTTATATATATTAACCTATAA 2937

GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H
; APPLICANT: Kumari, Kshama
; APPLICANT: DeAngelis, Paul
; TITLE OF INVENTION: STREPTOCOCCUS EQUITIMILIS HYALURONAN SYNTHASE GENE AND EXPRESS
; FILE REFERENCE: 3554.049
; CURRENT APPLICATION NUMBER: US/09/879,959
; CURRENT FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: 09/469,200
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 2937
; TYPE: DNA
; ORGANISM: pasteurella multocida
US-09-879-959-9

Query Match 76.6%; Score 2283.4; DB 33; Length 2937;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 2540; Conservative 0; Mismatches 376; Indels 21; Gaps 1;
QY 43 ATTTTAAAGGAAGAAATGAATATACATATACACAGCAATAAAGCATATAACGCAAT 102
Db 1 ATTTTAAAGGACAGAAATGAATATATATATATATATATATATATATATATATATAT 60
QY 103 GACTATGAATAGCACTCAATTTATTTGAGAGCTCTCTGAAACCTACGGCGAAATC 162
Db 61 GACTATCAATAGCACTCAATTTATTTGAAAGTCGGCGAAATCTATGGACGAAAT 120
QY 163 GTTGAATCCAAATTTATCAATGTAAAGAAACCTCTCGACCAATC----- 209
Db 121 GTTGAATTTCAATTTACCAATGCCAAGAAACCTCTCAGCACATCTCTCTGTTAATTC 180
QY 210 -----TTATGTAAGTGAAGATAAATAAAGAGTGTTCGCGATAGCTCAATATATC 261
Db 181 GCACATCTTCTGTAAATAAAGAAAGTCAATGTTTCGCGATAGTTCGTTAGATAT 240
QY 262 GCACACAGCTTCTTACCTTCCACAGTAAATAAATAAATTAATCTCTCGAATCAGAAAAAC 321
Db 241 GCACACAGCTTCTTACCTTCCACAGTAAATAAATAAATTAATCTCTGACTCGAAAAAC 300
QY 322 AGTTTAAAAATAAATGAAATCTATCACTGGGAAAAATCGGAGAACGAGAAATCAGA 381
Db 301 ACGTTAAAAATAAATGAAATCTCTCACTGAGAAGAAATCTGAAATCGGAGGTAAGA 360
QY 382 AAGTGAACACTAGTACCACAAAGATTTCCCTAAAGATCTGTTCTGCTCCATTTGCCAGAT 441
Db 361 CGGTCGCCCTTGTACCAAAAGATTTTCCCAAGATCTGGTTTACGCCCTTTACCTGAT 420
QY 442 CATGTTAATGATTTTACATGTCACAAAAATCGAAAAAAGCTTAGGTATATAAGCTGTA 501
Db 421 CATGTTAATGATTTTACATGTCACAAAAAGCAAGAAAGAGCTTGGCATATAAAGCTGAA 480
QY 502 AATAAGAAATCGGCTCTTCTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 561
Db 481 CATCAACATGTTGGTCTTCTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 540
QY 562 ATAAGTTAGCTGTTTGGTCAATCAGAAAAACAACTACCCATTTGAAGTCTGTTGCA 621
Db 541 ATTACATTAGCTGTTTGGTCAATCAGAAAAACAACTACCCATTTGAAGTCTGTTGCA 600
QY 622 GATGTTAGTGAAGAAACCTTACTTACCAATGTCAAAAATACGACAAAACTTGAC 681
Db 601 GATGTTAGTGAAGAAACCTTACTTACCAATGTCAAAAATACGACAAAACTTGAC 660
QY 682 ATAAAGTATGAAGACAAAAAGATTTATGATATCAATTTGTGTGAGTCAAGAACTAGCT 741
Db 661 ATTCGCTACGTACAGAAAAAGATAACGGTTTTCAGCGAGTGGCGCTCGGAATATGGGA 720

Qy 742 TTACGTACAGCAAGATGATGATTTGCTCGATTCGTAGCTGCGATATGCGACCAACA 801
Dy 721 TTACGCTTAGCAAAATGACTTTATTTGCTTACTCGACTGTGATATGCGCCAAATCCA 780
Qy 802 TTATGGGTTTCATTTTATCTTACAGAACTATTTAGAGACAATGATATGTTTAAATTTGA 861
Dy 781 TTATGGGTTTCATTTTATGTTGACAGACTATTTAGAGAGATGATATTAACATCATCTTGT 840
Qy 862 CTTAGAAAATGTTGGATGATCAATATATACCGCAGAACAAATTCCTTAACGCATCATAT 921
Dy 841 CCAAGAAATACATGATACACAATATTTGACCCAAAGAGCTTCTTAATTAACCGGAGT 900
Qy 922 TTAATPAGAACTACCTGAAACCGCTACAAATAACAATCTCTGATTAACATCAAAAGGA 981
Dy 901 TTGCTTGAATCATTTACAGAGTGAACCAATAATAGTTGCGCCAAAGGGGAAGGA 960
Qy 982 AATATATCGTTGGATGAGATGAGAACTTTCAAAAAACCGATAATCTAGCTCTATGT 1041
Dy 961 ACAGTTTCTCGATTTGGCGCTTAGAACATTTGGAAGAAACAGAAAATCTCCGCTTATCC 1020
Qy 1042 GATTCCTCGTTTCTGTTATTTGTTGCGGCTAATGTTGCTTTTCTAAAGAAATGGCTAAAT 1101
Dy 1021 GATTCGCTTTCCGTTTTCGCGGGGTAAATGTTGCTTTCGCTTAAAGAAATGGCTAAAT 1080
Qy 1102 AAAGTAGGTTGTTGCTGATCAAGAAATTTAATCATTTGGGGGGGGAAGATGTAAGAAATTTGTT 1161
Dy 1081 AAATCCGCTTCTTTGATGAGGAATTTAATCACTGGGTGGAGAGATGCGAAATTTGGA 1140
Qy 1162 TACAGATATTTGCGCAAGGCTGTTTTCAGAGTAATTTAGCGGGGAATGGCCATCCAT 1221
Dy 1141 TATCGCTTATTTCCGTTACGGTAGTTCTTTTAAACATATTTGATGGCATTTATGGCTTACCAT 1200
Qy 1222 CAAGAACCACTGGTAAAGAAATCAACAGAACCGGAAGCTGGTAAAGATTTAGGCTT 1281
Dy 1201 CAAGAGCCACCGAGTAAAGAAATGAAACCGATCTGAAGCGGGGAAGAAATATTTACGCTC 1260
Qy 1282 AAAATTTGGAAGAAAGGTACCTTACATCTATAGAAAGCTTTTACCATAAGAAATTTCA 1341
Dy 1261 GATATTTAGAGAAAGGTCCCTTATATCTATAGAAATTTTACCATAAGAAATTTGG 1320
Qy 1342 CATATTTAGAAATACCTTTTATGTTCTATTTATATCCCGCTTATAACTGTGCAATTTAT 1401
Dy 1321 CATATCAATAGAGTACCTTTTATGTTTCAATTTATATCCAGCTTATAACTGTGCAAACTAT 1380
Qy 1402 ATTCAGAGTGTAGATAGTCTCTTAAATCAAACTGTTGTCGATCTCGAGGTTGTTAT 1461
Dy 1381 ATTCAGGTTGGTAGATAGTCACTGAATCAGATGTTGTTGTTGTTGTTGTTGTTAT 1440
Qy 1462 TGTACGATGTTTCAACAGATAATACCTTAGAAGTATCAATTAAGCTTTTATGGTAAATAT 1521
Dy 1441 TGTACGATGTTTCAACAGATAATACCTTAGAAGTATCAATTAAGCTTTTATGGTAAAT 1500
Qy 1522 CTTAGGTTACGATCATGTTTAAACCAATGCGGGAATAGCCCTCAGCATCAATGCGAGCC 1581
Dy 1501 CTTAGGTTACGATCATGTTTAAACCAATGCGGGAATAGCCCTCAGCATCAATGCGAGCC 1560
Qy 1582 GTTCTTTTGTCTTAAAGGTTATTTACATTTGGGAGTTAGATTTAGATTTATCTTGAGCCT 1641
Dy 1561 GTTCTTTTGTCTTAAAGGTTATTTACATTTGGGAGTTAGATTTAGATTTATCTTGAGCCT 1620
Qy 1642 GATGAGTTGACTGTGTTTAAAGAAATTTTAAAGATTAACACGCTAGCTTTGTTGTTAT 1701
Dy 1621 GATGAGTTGACTGTGTTTAAAGAAATTTTAAAGATTAACACGCTAGCTTTGTTGTTAT 1680
Qy 1702 ACCACTAATAGAAAGCTCAATCCGATGGTAGCTTAAATGCGTAAATGGTTTACAAATTTGGA 1761
Dy 1681 ACCACTAATAGAAAGCTCAATCCGATGGTAGCTTAAATGCGTAAATGGTTTACAAATTTGGA 1740
Qy 1762 GAATTTTACAGAGAAACTCACACGCTATGATTTGCTCACCATTATAGATGTTTACG 1821
Dy 1741 GAATTTTACAGAGAAACTCACACGCTATGATTTGCTCACCATTATAGATGTTTACG 1800
Qy 1822 ATTAGAGCTTGGCATTTTAAACGATGGATTTTAAACGAAATATTTGAAACCGCTGGATTTAT 1881

Dy 1801 ATTAGAGCTTGGCATTTAACTGATGATCAATGAAGAAATTTCAAAATGCGGTAGACTAT 1860
Qy 1882 GACATGTTCTTAACTCAGTGAAGTTGAAATTTTAAACATCTTTAAATTAATCTGCTAT 1941
Dy 1861 GACATGTTCTTCAAACTCAGTGAAGTTGAAATTTTAAACATCTTTAAATTAATCTGCTAT 1920
Qy 1942 AACCGGTTATTTACATGTTGATTAACACATCAATTAAGAACTGCGATTTCAAAAGAAAC 2001
Dy 1921 AACCGGTTATTTACATGTTGATTAACACATCAATTAAGAACTTGGCATTTCAAAAGAAAC 1980
Qy 2002 CATTTTGTGTAGTCAATCAGTCAATTTAAATPAGACAGGATCAATTTATTAATTTATGAC 2061
Dy 1981 CATTTTGTGTAGTCAATCAGTCAATTTAAATPAGACAGGATACTTTATTAATTTATGAC 2040
Qy 2062 AAATTTGATGATTTAGATGAAGTACAAAGTATATCTTCAATTAACCGCTGATATCAA 2121
Dy 2041 GAATTTGATGATTTAGATGAAGTACAAAGTATATTTTCAATTAACCGCTGATATCAA 2100
Qy 2122 GAAGAAATGGATGATTTAAAGATCTTAACTCAATTAAGATGCAAAATCGCA 2181
Dy 2101 GAAGAGATTTGATATCTTAAAGATATTTAAATCATCCAGATTAAGATGCAAAATCGCA 2160
Qy 2182 GTCAGTATTTTCTATCCCAATACATTTAAACGGCTTAGTGAAGAACTTAAACATTTATTT 2241
Dy 2161 GTCAGTATTTTCTATCCCAATACATTTAAACGGCTTAGTGAAGAACTTAAACATTTATTT 2220
Qy 2242 GAATTAATTAAGAAATATTTTCTGTTATTTTCTACATGTTGATTAAGAACTATCTTACACCA 2301
Dy 2221 GAATTAATTAAGAAATATTTTCTGTTATTTTCTACATGTTGATTAAGAACTATCTTACACCA 2280
Qy 2302 GACATCAAAAGAAATATTTGGCTTCTATCAATGAACCAAGTGAATTTTACTAAAT 2361
Dy 2281 GATATCAAAAGAAATATTTAGCTTCTATCAATGAACCAAGTGAATTTTACTAAAT 2340
Qy 2362 AATGACATCTCATATTTACAGAGTAATAGACTAATTAAGAACTGAGGACATCTTAAAGTAAT 2421
Dy 2341 AATGATATCTCATATTTACAGAGTAATAGACTAATTAAGAACTGAGGACATCTTAAAGTAAT 2400
Qy 2422 ATTAATTAATTAAGTCAAGTAAATCTAATTTGTAATACATCATTTTGTATATCATGAC 2481
Dy 2401 ATTAATTAATTAAGTCAAGTAAATCTAATTTGTAATACATCATTTTGTATATCATGAC 2460
Qy 2482 AGCTATTTGTTTAAAGTACAGCTATGCTTATATGAAGAAATATGATGTCGACATCAAT 2541
Dy 2461 AGCTATTTGTTTAAAGTACAGCTATGCTTATATGAAGAAATATGATGTCGACATCAAT 2520
Qy 2542 TTCTCAGCATTAACACATGATTGGATCGAGAAATCAATGCGCATCCACCATTTAAAG 2601
Dy 2521 TTCTCAGCATTAACACATGATTGGATCGAGAAATCAATGCGCATCCACCATTTAAAG 2580
Qy 2602 CTGATTTAAACCTATTTTAAATGACATGACTTAAAGATGATGAATGTAAGAGGATCA 2661
Dy 2581 CTGATTTAAACCTATTTTAAATGACATGACTTAAAGATGATGAATGTAAGAGGATCA 2640
Qy 2662 CAAGGTATGTTTAAAGTATGCGCTACCGCATGAGCTTCTGACGATTTTAAAGAGATC 2721
Dy 2641 CAAGGTATGTTTAAAGTATGCGCTACCGCATGAGCTTCTGACGATTTTAAAGAGATC 2700
Qy 2722 ATCAGATCTGCAATCAATTTGATGTTGCCAGATATTAACACTGAGGATTTTGGTTC 2781
Dy 2701 ATCAGATCTGCAATCAATTTGATGTTGCCAGATATTAACACTGAGGATTTTGGTTC 2760
Qy 2782 CAATTTGCACTTTTAAATCTTAGAAAAGAAACCGGCATGATTTTAAATTAACATCGACC 2841
Dy 2761 CAATTTGCACTTTTAAATCTTAGAAAAGAAACCGGCATGATTTTAAATTAACATCGACC 2820
Qy 2842 CTGACTTATATGCTTTGGGAACGAAATTTCAATGGAACAAATGAACAAATTTCAAGTGCA 2901
Dy 2821 CTGACTTATATGCTTTGGGAACGAAATTTCAATGGAACAAATGAACAAATTTGAAGTGCA 2880
Qy 2902 AAAAAAGGGGAAATATCCCGGTTAAACAGTTTCAATTTATTAATAGTATAACGCTATAA 2958
Dy 2880 AAAAAAGGGGAAATATCCCGGTTAAACAGTTTCAATTTATTAATAGTATAACGCTATAA 2958

Db 2881 AAAAGAGGAGAAATATACCTGTTAAACAAGTTCATTATTAAAGTAGTAACTATATA 2937

RESULT 10

US-10-172-527-9

; Sequence 9, Application US/10172527

; GENERAL INFORMATION:

; APPLICANT: Weigel, Paul H

; APPLICANT: Kumari, Kshama

; APPLICANT: DeAngelis, Paul

; TITLE OF INVENTION: HYALURONAN SYNTHASE GENES AND EXPRESSION THEREOF IN BACILLUS SUBT

; FILE REFERENCE: 3554.048

; CURRENT APPLICATION NUMBER: US/10/172,527

; CURRENT FILING DATE: 2002-06-13

; PRIOR APPLICATION NUMBER: 60/297,788

; PRIOR FILING DATE: 2001-06-13

; PRIOR APPLICATION NUMBER: 60/297,744

; PRIOR FILING DATE: 2001-06-13

; PRIOR APPLICATION NUMBER: 09/469,200

; PRIOR FILING DATE: 1999-12-21

; PRIOR APPLICATION NUMBER: 09/178,851

; PRIOR FILING DATE: 1998-10-26

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 9

; LENGTH: 2937

; TYPE: DNA

; ORGANISM: pasteurella multocida

; US-10-172-527-9

Query Match 76.6%; Score 2283.4; DB 41; Length 2937;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 2540; Conservative 0; Mismatches 376; Indels 21; Gaps 1;

Qy 43 ATTTTAAAGGAAGAAATGAATACATTATCACAGCAATATAAGCATATAACAGCAAT 102
Db 1 ATTTTAAAGGACAGAAATGAATACATTATCACAGCAATATAAGCATATAACAGCAAT 60
Qy 103 GACTATGAATAGCACACAAATATTGAGAAGTCTGCTGAAACCTACGGCGGAAAAATC 162
Db 61 GACTATCAATAGCACACAAATATTGAAAAGTCGGCGGAAATCTATGGACGGAAT 120
Qy 163 GTTGAATTCAAATATCAATATGAAGAAACCTCTCGACCAATC----- 209
Db 121 GTTGAATTCAAATATCAATATGAAGAAACCTCTCGACCAATC----- 180
Qy 210 -----TTATGTAAGTGAAGATAAAAAACACAGTGTTCGGATAGCTCATTTAGATATC 261
Db 181 GCACATCTTTCTGTAATATAAGAAAGAAAGTCAATGTTGCGATAGTCCGTTAGATAT 240
Qy 262 GCAACACAGCTCTTACTTTCCAGGTAAAAAATAAATCTATCCGAAATCAGAAAAAAC 321
Db 241 GCAACACAGCTTACTTTCCAGGTAAAAAATAAATCTATCTGACTCGGAAAAAAC 300
Qy 322 AGTTTAAAAATAAATGAATATCTATCACTGGGAAAAAATCGGAGAACGCGAAATCAGA 381
Db 301 ACGTTAAAAATAAATGAATATCTGCTACTGAGAGAAATCTGAAATCGGAGGTAAGA 360
Qy 382 AAGTGGAACTAGTACCCAAAGATTTTCCCTAAAGATCTGTTGCTGCTCCATTGCGAGAT 441
Db 361 GCGGTGCGCCCTGTATACCAAGATTTTCCCAAGATCTGTTTACGCGCTTACCTGAT 420
Qy 442 CATGTTAATGATTTTACATGTTTACAAAAATCGAAAAAAGCTTAGGTATATAAGCCTGTA 501
Db 421 CATGTTAATGATTTTACATGTTTACAAAAAGCGGAAAAAGACTTGGCATAAAACTGAA 480
Qy 502 AATAGAAATATCGGCTCTTCTATTTATTTATTCCTACATTTAATCGTAGCGGTATTTAGAT 561
Db 481 CATCAACATGTGGTCTTCTATTTATTCGTTTACAAATCAATCGACACGAAATTTATCG 540
Qy 562 ATAACGTTAGCCGTTTGGTCAATCAGAAAAACAACTACCCCAATTTGAAGTCGTTTGA 621
Db 541 ATTACATTAGCCCTGTTTAGTAACCAACCAAAACACATATACCCGTTTGAAGTTATCGTGACA 600

Qy 622 GATGATGGTAGTAAGGAAAACTTACTTACCAATTTGTCAAAAAATACGAACAAAACTTGAC 681
Db 601 GATGATGGTAGTCAGGAAGATCTATCACCAGTATCATCCCAATATGAAATAAATTTGAT 660
Qy 682 ATAAAGTATGTAAGACAAAAAGATTATGGATATCAATTTGTGCAGTCAGAACTTAGT 741
Db 661 ATTCGCTACGTCAGACAAAAAGATAACGGTTTTCAGCCAGTCCGCTCGGAATATGGGA 720
Qy 742 TTACGTCACGCAAAAGTATGATTTTGTCTCGATTCTAGACTCGCATATGGCACCACAA 801
Db 721 TTACGCTTACAAAATATGACTTTTGGCTTACTCCACTGTGATATGGCCCAAAATCCA 780
Qy 802 TTATGGGTTTCATTTCTTATCTTACAGAACTATTTAGAACAAATGATATTTTAAATGGA 861
Db 781 TTATGGGTTTCATTTCTTATGTCAGAGCTATTTAGAAATGATGATTTTAAACAATCATGGT 840
Qy 862 CCTAGAAAAATGTGGATCTCATATAATTTACCCGACAGCAATTTCTTAAAGGATCCATAT 921
Db 841 CCAAGAAAAATACATCGATACACAACTATTTGACCCAAAAGACTTCTTAAATAACCGGAGT 900
Qy 922 TTAATAGAATCACTACCTGAAACCGCTACAAATAACCAATCTTCGATTACATCAAAAGGA 981
Db 901 TTGCTTGAATCATTTACCAGAGTGAACCAATAATAGTGTTCGCGCAAAAGGGGAAGGA 960
Qy 982 AATATATCGTTGGATTTGAGATTTAGAACATTTCAAAAAACCGGATATCTAGCTATATGT 1041
Db 961 ACAGTTTCTCTGGATTTGGCGCTTAGAACAAATTCGAAAAACAGAAAAATCTCCGCTTATCC 1020
Qy 1042 GATTCTCGGTTTCGTTATTTTGTTCGGGTAATGTTGCGATTCTTAAAGATGGCTAAAT 1101
Db 1021 GATTGCGCTTTCCGTTTTCGCGGGGTAATGTTGCTTAAAGATGGCTAAAT 1080
Qy 1102 AAAGTAGTGTTCGATGAAGAATTTAATCATTTGGGGGGGGAAGATGTAGATTTGGT 1161
Db 1081 AAATCGGTTTCTTTGATGAGGAATTTAATCACTGGGTGGAGAGATGTGGAATTTGGA 1140
Qy 1162 TACAGATTTATTTGCCAAGGCTGTTTTTTCAGAGTAATTTGACGGCGGATGGCCATCAT 1221
Db 1141 TATCGCTTATTTCCGTTACGGGTAGTTTCTTAAACTATTTGATGCGATTTATGGCCTACCAT 1200
Qy 1222 CAAGAACCCCTGTTGTAAGAAAAATGAAACAGAACCGAGCTGGTAAAAGTATTTAGCTTT 1281
Db 1201 CAAGAGCCACCCAGGTAAAGAAAAATGAAACCGATCTGGAACGGGAAAAAATATTACGCTC 1260
Qy 1282 AAAATGTGAAAGAAAGGTACCTTACATCTATAGAAAGCTTTTACCAATAGAGATTTCA 1341
Db 1261 GATATTTAGAGAAAAAGGTCCCTTATATCTATAGAAAACTTTTACCAATAGAGATTTG 1320
Qy 1342 CATATTTAGAAATACCTTTTACTTTTATATATCCCGCTTATAACTGTGCAAAATAT 1401
Db 1321 CATATCAATAGATACCTTTTACTTTTATATATATATATATATATATATATATATATAT 1380
Qy 1402 ATTTCAAGATGTTAGATAGTCTCTTAACTAAACTGTTTGTCTGATCTCGAGGTTTGTAT 1461
Db 1381 ATTTCAAGCTTGGTAGATAGTACCTGATCAGACTGTTTGTGATCTCGAGGTTTGTAT 1440
Qy 1462 TGTAAAGATGTTTACAGATTAATACCTTAGAAGTATCAATTAAGCTTTTATGGTATAT 1521
Db 1441 TGTAAAGATGTTTACAGATTAATACCTTAGAAGTATCAATTAAGCTTTTATGGTATAT 1500
Qy 1522 CCTAGGTCAGCATCATGTCTAAACCAATGCGGAATAGCCTTCAGCATCAATGACGCC 1581
Db 1501 CCTAGGTCAGCATCATGTCTAAACCAATGCGGAATAGCCTTCAGCATCAATGACGCC 1560
Qy 1582 GTTCTCTTTTGTAAAGGTTTATACATTTGGCAGTTTATAGATTCAGATGATTTCTTTAGCCT 1641
Db 1561 GTTCTCTTTTGTAAAGGTTTATACATTTGGCAGTTTATAGATTCAGATGATTTCTTTAGCCT 1620
Qy 1642 GATGAGTTGAACGTGTTTAAAGAAATTTTAAAGATAAAAACGCTAGCTCTGTTGTTAT 1701
Db 1621 GATGAGTTGAACGTGTTTAAAGAAATTTTAAAGATAAAAACGCTAGCTCTGTTGTTAT 1680

Qy 1702 ACCACTAATAGAAACGTCATCCGGATGGTACGCTTAATCGCTAATGGTTACAATGGCCA 1761
Db 1681 ACCACTAATAGAAACGTCATCCGGATGGTACGCTTAATCGCTAATGGTTACAATGGCCA 1740
Qy 1762 GAATTTTCACGAGAAAACTCACAACGGCTATGATGCTACCAATTTAGAAATGTTTACG 1821
Db 1741 GAATTTTCACGAGAAAACTCACAACGGCTATGATGCTACCAATTTAGAAATGTTTACG 1800
Qy 1822 ATTAGACTTGGCATTTAAACGGATGGATTTAAACGAAATATTTGAAAGCCGCTGATAT 1881
Db 1801 ATTAGACTTGGCATTTAAACGGATGGATTTAAACGAAATATTTGAAAGCCGCTGATAT 1860
Qy 1882 GACATGTTCCCTAAACTCAGTGAAGTTGGAAAAATTTAAACATCTTAATAAAATCTGCTAT 1941
Db 1861 GACATGTTCCCTAAACTCAGTGAAGTTGGAAAAATTTAAACATCTTAATAAAATCTGCTAT 1920
Qy 1942 AACCGGTATTTACATGTTGATGATAACACATCCATTAAGAACTCGGCATTCAAAAGAAAAAC 2001
Db 1921 AACCGGTATTTACATGTTGATGATAACACATCCATTAAGAACTCGGCATTCAAAAGAAAAAC 1980
Qy 2002 CATTTTGTGTGAGTCAATCAGTCAATTAATAGACAAGGCATCAATTTATTAATATATGAC 2061
Db 1981 CATTTTGTGTGAGTCAATCAGTCAATTAATAGACAAGGCATCAATTTATTAATATATGAC 2040
Qy 2062 AAATTTGATGATTTAGATGAAGTAGAAAGTATATCTTCAATAAAACCGCTGAATATCAA 2121
Db 2041 GAATTTGATGATTTAGATGAAGTAGAAAGTATATTTTCAATAAAACCGCTGAATATCAA 2100
Qy 2122 GAAGAAATGGATGTTTAAAGATCTTAAAGTCAATTAAGATGATGCAAAATCCGA 2181
Db 2101 GAAGAGATTTGATATCTTAAAGATATTTAAATATCATCCAGAAATAAGATGCCAAATCCGA 2160
Qy 2182 GTCAGTATTTTCTATCCCAATACATTAACGGCTTTAGTGAAGAACTAAACAAATATATT 2241
Db 2161 GTCAGTATTTTCTATCCCAATACATTAACGGCTTTAGTGAAGAACTAAACAAATATATT 2220
Qy 2242 GAATATATAAATAATATTCGTTATTATCTACATGTTGATAAGAAATCATCTTACACCA 2301
Db 2221 GAATATAAATAAATAATATTCGTTATTATCTACATGTTGATAAGAAATCATCTTACACCA 2280
Qy 2302 GACATCAAAAAGAAATATTTGGCTTTCTATCATTAAGCAGCAAGTGAATATTTTACTAAAT 2361
Db 2281 GATATCAAAAAGAAATATTTAGCTTTCTATCATTAACATCAAGTGAATATTTTACTAAAT 2340
Qy 2362 AATGACATCTCATATTACACAGTAATAGACTAATAAAGCTGAGGCATTTAAGTAAT 2421
Db 2341 AATGATATCTCATATTACACAGTAATAGACTAATAAAGCTGAGGCATTTAAGTAAT 2400
Qy 2422 ATTAATAAATTAAGTCAGTTAAATCTAAATTTGTAATACATCAATTTTGTAAATCATGAC 2481
Db 2401 ATTAATAAATTAAGTCAGTTAAATCTAAATTTGTAATACATCAATTTTGTAAATCATGAC 2460
Qy 2482 AGCCTATTCGTTAAATATGACAGCTATGCTTTATATGAAAAATATGATGTCGGCATGAAT 2541
Db 2461 AGCCTATTCGTTAAATATGACAGCTATGCTTTATATGAAAAATATGATGTCGGCATGAAT 2520
Qy 2542 TTCTCAGCATTAACACATGATTTGATCGAGAAATCAATGCGCATCCACCAATTTAAAAAG 2601
Db 2521 TTCTCAGCATTAACACATGATTTGATCGAGAAATCAATGCGCATCCACCAATTTAAAAAG 2580
Qy 2602 CTGATTTAAACCTATTTTAAATGACAATGACTTAAGAGATGATGAATGTGAAAGGGGCATCA 2661
Db 2581 CTGATTTAAACCTATTTTAAATGACAATGACTTAAGAGATGATGAATGTGAAAGGGGCATCA 2640
Qy 2662 CAAGGTATGTTTATGAAGTATGCGCTACCGCATGAGCTTCTGACCATTTATTAAGAAAGTC 2721
Db 2641 CAAGGTATGTTTATGAAGTATGCGCTACCGCATGAGCTTCTGACCATTTATTAAGAAAGTC 2700
Qy 2722 ATCACATCTGCCAGTCAATTTGATAGTGTGCCAGAAATAACACTGAGGATATTTGGTTC 2781
Db 2701 ATCACATCTGCCAGTCAATTTGATAGTGTGCCAGAAATAACACTGAGGATATTTGGTTC 2760
Qy 2782 CAATTTGCACCTTTTAAATCTTAGAAAAAGAAAAACCGGCCATGTATTTTAAATAAACATCGACC 2841

Db 2761 CAATTTGCACCTTTTAAATCTTAGAAAAAGAAACCCGCCATGTATTTTAAATAAACATCGACC 2820
Qy 2842 CTGACTTTATATGCTTTGGGAACGAAATTTACAATGGACAAATGAACAAATTTCAAAGTGCA 2901
Db 2821 CTGACTTTATATGCTTTGGGAACGAAATTTACAATGGACAAATGAACAAATTTCAAAGTGCA 2880
Qy 2902 AAAAAAGCGGAAATATCCCGTTAAACAGTTTCATTATTAATAGTATACGCTATATAA 2958
Db 2881 AAAAGAGGAGAAATATACCTGTTAAACAAGTTTCATTATTAATAGTATACCTATATAA 2937
RESULT 11
PCT-US99-26501-2
; Sequence 2, Application PC/TUS9926501
; GENERAL INFORMATION:
; APPLICANT: THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA
; TITLE OF INVENTION: POLYMER GRAFTING BY POLYSACCHARIDE SYNTHASES
; FILE REFERENCE: 617481-5
; CURRENT APPLICATION NUMBER: PCT/US99/26501
; CURRENT FILING DATE: 1999-11-10
; EARLIER APPLICATION NUMBER: 60/107,929
; EARLIER FILING DATE: 1998-11-11
; EARLIER APPLICATION NUMBER: 09/283,402
; EARLIER FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2112
; TYPE: DNA
; ORGANISM: Pasteurella multocida
PCT-US99-26501-2
Query Match 50.5%; Score 1503.2; DB 1; Length 2112;
Best Local Similarity 82.5%; Pred. No. 2.9e-273;
Matches 1743; Conservative 0; Mismatches 348; Indels 21; Gaps 1;
Qy 61 ATGAATACATTTATCAAGCAATAAAGCATATAACAGCAATGACTATGCAATTAGCACTC 120
Db 1 ATGAATACATTTATCAAGCAATAAAGCATATAACAGCAATGACTATCAATTAGCACTC 60
Qy 121 AAATTTATTTGGAAGTCTGCTGAAACCTACGGGGGAAAAAATCGTTGAAATTCAAATATTC 180
Db 61 AAATTTATTTGGAAGTCTGCGGGAATCTATGACGGAATAATGTTGAAATTCAAATATTC 120
Qy 181 AAATCTAAGAAAAACTCTCGACCAATTC-----TTATGTAAGT 219
Db 121 AAATGCCAAGAAAAACTCTCAGCACATCTTCTGTAATTCAGCACATCTTCTGTGTAAT 180
Qy 220 GAAGATAAAAAAAGACATGTTTTCGATAGCTCATTTAGATATCGCAACACAGCTCTTACTT 279
Db 181 AAGAAGAAAAAAGTCAATGTTTTCGATAGCTGTTGCAACACACATGTTACTT 240
Qy 280 TCCAAAGTAAAAAATTAATCTATCCGAATCAGAAAAAAGAGTTTAAAAAATAAATGG 339
Db 241 TCCAAAGTAAAAAATTTAGTACTTTCTGACTCGGAAAAAAGACGTTAAAAAATAAATGG 300
Qy 340 AAATCTATCACTGGGAAAAAATCGGAGACGAGAAATCAGAAAGTGAACCTAGTATACC 399
Db 301 AAATGCTCAGTGAAGAAAAATCTGAAAAATCGGAGGTAAAGCGGTGCCCTTGTACCA 360
Qy 400 AAAGATTTTCTTAAAGATCTTCTTCTGCTCCATTTGCCAGATCATGTTAATGATTTTACA 459
Db 361 AAAGATTTTCCCAAGATCTGCTTTTAGCGCTTTTACCTGATCATGTTAATGATTTTACA 420
Qy 460 TGGTACAAAAATCGAAAAAAGAGCTTAGGTATAAGCCCTGTAAATGAAGAAATATCGGCTTT 519
Db 421 TGGTACAAAAAGGAAAAAAGAGCTTAGGTATAAAGCTGGAACATCAACATGTTGGCTTT 480
Qy 520 TCTATTATTTCTTACATTTTAAATCGTACCGCTATTTTATAGATATACGTTAGCCCTGTTG 579
Db 481 TCTATTATCGTTACACATTTCAATCGACGCAATTTTATCATGATTACATTCGCTGTTTAA 540

QY 580 GTCAATCAGAAACAACTACCCATTGTAAGTCGTTGTTGCAGATGATGGTAGTAAGGAA 639
DB 541 GTAAACCAAAACACATTACCGCTTGAAGTTATCGTCACAGATGATGGTAGTCAGGAA 600
QY 640 AACTTACTTACCATTGTGCAAAATAACGAAACAAACTTGCATAAAGTAGTGTAGACAAA 699
DB 601 GATCTATCACCAGTATCTCCCAATATGAATAAATAATTCGATTCAGCTACGACAAA 660
QY 700 AAGATATATGATATCAATTTGTCGAGTCAGAACTTAGGTTTACGTACAGCAAAAGTAT 759
DB 661 AAGATAACGGTTTTCAGCCGAGTCGCGCTCGGAATATGGGATTAGCCTTAGCAAAATAT 720
QY 760 GATTTTGTCTCGATTCTAGACGTCGATATGGCACCAACAATAATATGGGTTTCATCTTAT 819
DB 721 GACTTTATTTGGCTTACTCGACTGTGATATGGCCCAATCCATATATGGGTTTCATCTTAT 780
QY 820 CTTACAGACTATTAGAAGCAATATGATTTGTTTAAATTTGGACCTAGAAAATATGTTGGAT 879
DB 781 GTTGCAGAGCTATTAGAAGATGATATTTAAACAATCATTTGGTCCAAAGAAATACATCGAT 840
QY 880 ACTCATAATATACCCGACAGCAATTTCTTAACGATCCATATTTAATAGAACTCACCTACCT 939
DB 841 ACACACATATTCACCAAAAGACTTCTTAATAACCGGAGTTTGGCTTGAATCATTTACCA 900
QY 940 GAAACCGCTACAATAACCAATCTCGATTTACATCAAAAAGGAAATATATCGTTGGATTTG 999
DB 901 GAAAGTAAACCAACATAATAGTTGTCGCGCAAAAGGGAAGGAAACAGTTTCTCTGGATTG 960
QY 1000 AGATTAGACATTTCAAAAACCGATATCTACGTCATGATCTCCGTTTCGTTTAT 1059
DB 961 CGCTTAGAACAAATTCGAAAACAGAAAATCTCCGCTTATCCGATTCGCGCTTTCCGTTT 1020
QY 1060 TTTGTTGCGGGTAATGTTGCTATTTCTTAAAGAATGCTTAAATAGGTAGTTGGTTTCGAT 1119
DB 1021 TTTGCGCGGGTAATGTTGCTTTCGCTTAAATAATGCTTAAATAATCCGTTTCTTTGAT 1080
QY 1120 GAAGAAATTTAATCATTTGGGGGCGGAAGATGTAGAATTTGGTTTACAGATTTATTTGCCAAA 1179
DB 1081 GAGGAATTTAATCACTTGGGTGGAGAAGATGTGGAATTTGGATATCGCTTATTTCCGTTAC 1140
QY 1180 GCGTGTGTTTTCAGAGTAATTTGACGGGGAATGGCCATCCATCAAGAACACCTGGTAAA 1239
DB 1141 GGTAGTTTCTTTAAACTATTTGATGGCTATTTGGCTTACCATCAAGAGCCACCGGTAAA 1200
QY 1240 GAAATGAAACAGACGAGCTGTGTAAGATTTACGCTTAAATTTGTGAAGAAAG 1299
DB 1201 GAAATGAAACCCGATCGTGAAGCGGGAATAATATACGCTGATATATGAGAGAAAG 1260
QY 1300 GTACCTTACATCTATAGAAAGCTTTTACCAATAGAGATTCACATATTCATAGAATACCT 1359
DB 1261 GTCCCTTATATCTATAGAAAACCTTTTACCAATAGAGATTCGATATCAATAGAGTACCT 1320
QY 1360 TTAGTTTCTATTTATATCCCGCTTATTAACCTGTGCAATTTATTTCAAGATGTGTAGAT 1419
DB 1321 TTAGTTTCAATTTATATCCCGCTTATTAACCTGTGCAATTTATTTCAACGTTGCGTAGAT 1380
QY 1420 AGTGCTCTTAATCAAACTGTTGTCGATCTCGAGTTTGTATTTGTAACGATGTTTCAACA 1479
DB 1381 AGTGCACTGAATCAGACTGTTGTTGATCTCGAGTTTGTATTTGTAACGATGTTTCAACA 1440
QY 1480 GATAATACCTTAGAAGTGATCAATAAGCTTATGTTAATAATCTAGGGTACGCATCATG 1539
DB 1441 GATAATACCTTAGAAGTGATCAATAAGCTTATGTTAATAATCTAGGGTACGCATCATG 1500
QY 1540 TCTAAACCAATGCGGAATAGCTCAGCATCAAAATCAGCCGTTTCTTTTGTCTAAAGGT 1599
DB 1501 TCTAAACCAATGCGGAATAGCTCAGCATCAAAATCAGCCGTTTCTTTTGTCTAAAGGT 1560
QY 1600 TATTACATTTGGCAGTTTAGATTCAGATGATATCTTTGAGCCTGATGAGTTGAACTGTGT 1659
DB 1561 TATTACATTTGGCAGTTTAGATTCAGATGATATCTTTGAGCCTGATGAGTTGAACTGTGT 1620
QY 1660 TTAAGAAGATTTTAAAGATAAACCCTAGCTGTGTTTATACCACTAATAGAAACGTC 1719

DB 1621 TTAAGAAGATTTTAAAGATAAACCCTAGCTGTGTTTATACCACTAATAAGAAAGTC 1680
QY 1720 AATCCGGATGGTAGCTTAATCGCTAATGTTACAATTTGGCCAGAAATTTTCCAGGAAAAA 1779
DB 1681 AATCCGGATGGTAGCTTAATCGCTAATGTTACAATTTGGCCAGAAATTTTCCAGGAAAAA 1740
QY 1780 CTCACACCGCTATGATTTGCTCACCATTTTTAGAATGTTTACGATTTAGAGCTTGGCATTTA 1839
DB 1741 CTCACACCGCTATGATTTGCTCACCATTTTTAGAATGTTTACGATTTAGAGCTTGGCATTTA 1800
QY 1840 ACGGATGGATTTAACGAAATATTTGAAACCGCGTGGATTTACATGTTTCCCTTAAACTC 1899
DB 1801 ACTGATGGATTTCAATGAAAAAATTTGAAATCGCGTAGACTATGACATGTTTCCCTCAAACTC 1860
QY 1900 AGTGAAGTTGGAAAAATTTTAAACATCTTAATAAATCTGTATTAACCCGCTATTACATGGT 1959
DB 1861 AGTGAAGTTGGAAAAATTTTAAACATCTTAATAAATCTGTATTAACCCGCTATTACATGGT 1920
QY 1960 GATAACACATCCATTAAAGAACTCGGCATTTCAAAAGAAAAACCATTTTGTTGTAGTCAAT 2019
DB 1921 GATAACACATCAATTAAGAAACTTGGCATTTCAAAAGAAAAACCATTTTGTTGTAGTCAAT 1980
QY 2020 CAGTCATTAAATAGACAGGCATCAATTTATATTAATTTATGCAATTTGATGATTAGAT 2079
DB 1981 CAGTCATTAAATAGACAGGCATCAATTTATATTAATTTATGCAATTTGATGATTAGAT 2040
QY 2080 GAAAGTAGAAGTATATCTTCAATAAACCGCTGAAATATCAAGAAGAAATGGATATGTTA 2139
DB 2041 GAAAGTAGAAGTATATCTTCAATAAACCGCTGAAATATCAAGAAGAAATGGATATGTTA 2100
QY 2140 AAAGATCTTAAA 2151
DB 2101 AAAGATATTAAA 2112

RESULT 12
US-09-107-532-2527
; Sequence 2527, Application US/09107532
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7308
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER:
; OPERATING SYSTEM:
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085598
; FILING DATE: May 14, 1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 2527:

SEQUENCE CHARACTERISTICS:
LENGTH: 984 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...984
US-09-107-532-2527

Query Match 3.0%; Score 89.2; DB 15; Length 984;
Best Local Similarity 54.2%; Pred. No. 1.3e-06;
Matches 181; Conservative 0; Mismatches 153; Indels 0; Gaps 0;
QY 1349 ATAGAATACCTTTAGTTCTTATTTATATCCCGCTTATTAACCTGTGCAAAATATATTCAAA 1408
Db ATATAATGTGAGATTAGTATTATTGCTGTGTATAAAGTTGAAAAATATTTAAGAA 73
QY 1409 GATGTGTAGTAGTCTCTTAATCAAACTCTTGTGATCTCGAGTTTGTATTGTTAACG 1468
Db AATGTGTAGATTGATTTAGCCCAACATTTACTGATTTTGAAGTTATATTAGTAGATG 133
QY 1469 ATGGTTCAACAGATATACCTTTAGAAGTGATCAATAGCTTTTATGTTAAATCCTTAGG 1528
Db ATGCATCTCTCGAATAGCGGGAATATCTGTGATGATGATGCTGTAAGAAAGATAATCGAG 193
QY 1529 TAGCGATCATGTCTAAACCAATGGCGGAATAGCCTCAGCATCAATGCGAGCGTTTCTT 1588
Db TAGTGTGATCCATTAAGAAACGGCGGATTAAGCAGTGTAGAAATGCGGGGATGATG 253
QY 1589 TTGCTAAAGGTTATTACATTTGGCAGTTAGATTTCAGATGATTATCTTTGAGCCTGATCGAG 1648
Db TTGCGAGAGCAAACTTAGGCTTTTGTGACAGTGATGATTATATTGATGAAGATATGT 313
QY 1649 TTGAACGTGTTTAAAGAAATTTTAAAGATTA 1682
Db ATGAAATTTTGTATGAGAATTTGAAATTCATGA 347

RESULT 13

US-09-107-532A-1527

Sequence 2527 Application US/09107532A
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107.532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 2527:
SEQUENCE CHARACTERISTICS:
LENGTH: 984 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...984
SEQUENCE DESCRIPTION: SEQ ID NO: 2527:
US-09-107-532A-2527

Query Match 3.0%; Score 89.2; DB 15; Length 984;
Best Local Similarity 54.2%; Pred. No. 1.3e-06;
Matches 181; Conservative 0; Mismatches 153; Indels 0; Gaps 0;
QY 1349 ATAGAATACCTTTAGTTCTTATTTATATCCCGCTTATTAACCTGTGCAAAATATATTCAAA 1408
Db ATATAATGTGAGATTAGTATTATTGCTGTGTATAAAGTTGAAAAATATTTAAGAA 73
QY 1409 GATGTGTAGTAGTCTCTTAATCAAACTCTTGTGATCTCGAGTTTGTATTGTTAACG 1468
Db AATGTGTAGATTGATTTAGCCCAACATTTACTGATTTTGAAGTTATATTAGTAGATG 133
QY 1469 ATGGTTCAACAGATATACCTTTAGAAGTGATCAATAGCTTTTATGTTAAATCCTTAGG 1528
Db ATGCATCTCTCGAATAGCGGGAATATCTGTGATGATGATGCTGTAAGAAAGATAATCGAG 193
QY 1529 TAGCGATCATGTCTAAACCAATGGCGGAATAGCCTCAGCATCAATGCGAGCGTTTCTT 1588
Db TAGTGTGATCCATTAAGAAACGGCGGATTAAGCAGTGTAGAAATGCGGGGATGATG 253
QY 1589 TTGCTAAAGGTTATTACATTTGGCAGTTAGATTTCAGATGATTATCTTTGAGCCTGATCGAG 1648
Db TTGCGAGAGCAAACTTAGGCTTTTGTGACAGTGATGATTATATTGATGAAGATATGT 313
QY 1649 TTGAACGTGTTTAAAGAAATTTTAAAGATTA 1682
Db ATGAAATTTTGTATGAGAATTTGAAATTCATGA 347

RESULT 14

US-09-134-000-2987
Sequence 2987, Application US/09134000A
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS
FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-005
CURRENT APPLICATION NUMBER: US/09/134,000A
CURRENT FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 6810
SEQ ID NO 2987
LENGTH: 993
TYPE: DNA
ORGANISM: Enterococcus faecalis
US-09-134-000-2987

Query Match 2.9%; Score 86; DB 15; Length 993;
Best Local Similarity 54.0%; Pred. NO. 5.3e-06;
Matches 176; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

[illegible]

RESULT 15
 US-09-070-927-137
 Sequence 137, Application US/09070927
 GENERAL INFORMATION:
 APPLICANT: Charles Kunsch
 APPLICANT: Patrick J. Dillon
 APPLICANT: Steven C. Barash
 TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and
 TITLE OF INVENTION: Polypeptides
 NUMBER OF SEQUENCES: 982
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/070,927
 FILING DATE: herewith
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: A. Anders Brookes
 REGISTRATION NUMBER: 36,373
 REFERENCE/DOCKET NUMBER: PB369
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 137:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5550 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-09-070-927-137

Qy	1354	ATACCTTTTAGTTTCTATTATATCCCGCTATAACTGTGCAAAATTATATTTCAAGAATGT	1413
Db	1001	ATCCCCAAAATTAGTATTATTTCTTCTGTATACAAATGTAGAAAAATATTTAGAAAAATGT	1060
Qy	1414	GTAGATAGTCCTCTTAATCAAACTGTGTCGAATCTCGAGGTTTGTATTTTGTAACGATGGT	1473
Db	1061	GTAGCCTCTATTTTACTCTCAACGTTTACTGACTTTGAATTCTCGTGCAGCATGGC	1120
Qy	1474	TCAACAGATAAATACCTTTAGAAGTGATCAATAAGCTTTTATGGTAAATAATCCTTAGGTCAGC	1533
Db	1121	TCTCCAGACAGAGTTCTCGGACCAATGTGTGATCAGTTTGTCTGAACAGAGATCAACGGGTAAA	1180
Qy	1534	ATCATGTCTAAACCAATGCGCGAATAGCGCTCAGCATCAAAATCGACGCGTTTCTTTTGCT	1593
Db	1181	GTTATCATAAACAAATGTTGGCTTAACCGATGCTCGTAAATGCTGGAATTTGAAATAGCA	1240
Qy	1594	AAAGGTTATTACATTGGGCAGTTAGATTTCCAGATGATTAATCTTTAGCCCTGATCGAGTTGAA	1653
Db	1241	ACAGGTGAGTATTTTAGGTTTCGTGATAGTATGATTAATCATTTGCAGATGATATGTATGAA	1300
Qy	1654	CTGTGTTTAAAGAATTTTTAAAGA	1679
Db	1301	CTATTATATCAAAATATCGTTAAAGA	1326

Search completed: January 4, 2003, 02:13:46
Job time : 5792 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen, Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2003, 21:32:29 ; Search time 249 Seconds
(without alignments)
9950.512 Million cell updates/sec

Title: US-09-842-484A-1
Perfect score: 2979
Sequence: 1 ttataactgattaaagaag.....acatttgctttattataaa 2979

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 657839 seqs, 415857249 residues

Total number of hits satisfying chosen parameters: 1315678

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_New.*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2979	100.0	2979	5	US-09-842-484A-1
2	2963	99.5	2979	5	US-09-842-484A-3
3	2285	76.7	2937	5	US-09-469-2000-8
4	2283.4	76.6	2937	6	US-10-309-560-7
5	2271.2	76.2	2920	6	US-10-217-613-2
6	814.2	27.3	14483	6	US-10-216-289-3
7	807.2	27.1	2058	6	US-10-216-289-1
8	86	2.9	993	5	US-09-134-000C-2987
c 9	79.2	2.7	11474	6	US-10-303-161-1
c 10	79.2	2.7	11474	6	US-10-303-162-1
c 11	79.2	2.7	11474	6	US-10-303-118-1
c 12	79.2	2.7	11474	6	US-10-303-128-1
c 13	79.2	2.7	11474	6	US-10-303-134-1
c 14	77.2	2.6	17527	6	US-10-240-454-27
c 15	76.6	2.6	13825	6	US-10-264-213-27
c 16	74.2	2.5	17276	5	US-09-751-708A-83
17	72.6	2.4	2529	5	US-09-134-000C-2344
18	72.4	2.4	25020	6	US-10-192-280-1
c 19	71.8	2.4	9106	6	US-10-240-452-69
c 20	70	2.3	14924	6	US-10-240-452-21
c 21	69.8	2.3	16258	6	US-10-257-166-120
c 22	68.4	2.3	891	6	US-10-303-161-30
23	68.4	2.3	891	6	US-10-303-162-30
24	68.4	2.3	891	6	US-10-303-118-30
25	68.4	2.3	891	6	US-10-303-128-30
26	68.4	2.3	891	6	US-10-303-134-30

ALIGNMENTS

RESULT 1

US-09-842-484A-1
; Sequence 1, Application US/09842484A
; GENERAL INFORMATION:
; APPLICANT: DEANGELIS, PAUL L.
; TITLE OF INVENTION: CHONDROITIN SYNTHASE GENE AND METHODS OF MAKING AND USING SAME
; FILE REFERENCE: 4605.003
; CURRENT APPLICATION NUMBER: US/09/842.484A
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2979
; TYPE: DNA
; ORGANISM: Pasteurella multocida
US-09-842-484A-1

Query Match		100.0%	Score 2979;	DB 5;	Length 2979;
Best Local Similarity		100.0%	Pred. No. 0;		
Matches 2979;		Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
Qy	1	TTATAACTGATTAAAGAAGGTAAACGATTCAAGCAAGGTTAATTTTAAAGGAAGAAA	60		
Db	1	TTATAACTGATTAAAGAAGGTAAACGATTCAAGCAAGGTTAATTTTAAAGGAAGAAA	60		
Qy	61	ATGATACATTATCACAGCAATAAAGCATATACAGCAATGACTATCAATTAGCACC	120		
Db	61	ATGATACATTATCACAGCAATAAAGCATATACAGCAATGACTATCAATTAGCACC	120		
Qy	121	AAATTTATTTGAGAACTCTCGAACCTACGGCGGCAAAAATCGTTGAATTCACAAATATC	180		
Db	121	AAATTTATTTGAGAACTCTCGAACCTACGGCGGCAAAAATCGTTGAATTCACAAATATC	180		
Qy	181	AAATTTAAAGAAAACTCTCGAACCTACGGCGGCAAAAATCGTTGAATTCACAAATATC	240		
Db	181	AAATTTAAAGAAAACTCTCGAACCTACGGCGGCAAAAATCGTTGAATTCACAAATATC	240		
Qy	241	TCCGATAGCTCTATTAGATATCCCAACACAGCTCTTACTTCCCAACGTAATAAATTAAC	300		
Db	241	TCCGATAGCTCTATTAGATATCCCAACACAGCTCTTACTTCCCAACGTAATAAATTAAC	300		
Qy	301	CTATCCGAATCAGAAAAAAGAGTTTAAAAATAAATGAAATCTATCACTCGGAAAAAA	360		
Db	301	CTATCCGAATCAGAAAAAAGAGTTTAAAAATAAATGAAATCTATCACTCGGAAAAAA	360		
Qy	361	TCGGAGAACGCAAAATCAGAAAGGTGGAACCTAGTACCCAAAGATTTTCTCTAAAGATCT	420		
Db	361	TCGGAGAACGCAAAATCAGAAAGGTGGAACCTAGTACCCAAAGATTTTCTCTAAAGATCT	420		

Sequence 162, Appl
Sequence 28, Appl
Sequence 28, Appl
Sequence 28, Appl
Sequence 28, Appl
Sequence 26, Appl
Sequence 26, Appl
Sequence 26, Appl
Sequence 26, Appl
Sequence 114, Appl
Sequence 18, Appl
Sequence 53, Appl
Sequence 138, Appl
Sequence 113, Appl
Sequence 148, Appl
Sequence 163, Appl
Sequence 36, Appl

Db 361 TCGGAGAACGAGAAATCAGAAAGCTGGNACTAGTACCCAGAGATTTTCCTAAAGATCTT 420
Qy 421 GTTCTTGCTCCATGCCAGATCATGTTAATGATTTTACATGGTACAAAATCGAAAAAA 480
Db 421 GTTCTTGCTCCATGCCAGATCATGTTAATGATTTTACATGGTACAAAATCGAAAAAA 480
Qy 481 AGCTTAGGTATAAAGCGCTGAATAGAAATATCGGCTCTTCTATTATTATCTCTACATTT 540
Db 481 AGCTTAGGTATAAAGCGCTGAATAGAAATATCGGCTCTTCTATTATTATCTCTACATTT 540
Qy 541 AATCGTAGCGGTATTTAGATATAAGCTTAGCCTGTTGGTCAATCAGAAAAACAACATAC 600
Db 541 AATCGTAGCGGTATTTAGATATAAGCTTAGCCTGTTGGTCAATCAGAAAAACAACATAC 600
Qy 601 CCATTTGAAGTCTGTTGCCAGATGATGTAGTAAGGAAAACTTACTTACCATTGTGCAA 660
Db 601 CCATTTGAAGTCTGTTGCCAGATGATGTAGTAAGGAAAACTTACTTACCATTGTGCAA 660
Qy 661 AATAGGAACAAAACCTTCACATAAAGTATGTAAGACAAAAAGATTTATGGATATCAATTG 720
Db 661 AATAGGAACAAAACCTTCACATAAAGTATGTAAGACAAAAAGATTTATGGATATCAATTG 720
Qy 721 TGTGAGTCAGAACTTAGTCTTACGTACAGCAAAAGTATGATTTGCTCGATCTAGAC 780
Db 721 TGTGAGTCAGAACTTAGTCTTACGTACAGCAAAAGTATGATTTGCTCGATCTAGAC 780
Qy 781 TGGGATATGGCACCACAAATATGGGTTCACTTCTTACAGAACTATTAGAAGAC 840
Db 781 TGGGATATGGCACCACAAATATGGGTTCACTTCTTACAGAACTATTAGAAGAC 840
Qy 841 AATGATATGTTTAAATTTGACCTAGAAAAATATGTTGGATCTCATATATACCGCAGAA 900
Db 841 AATGATATGTTTAAATTTGACCTAGAAAAATATGTTGGATCTCATATATACCGCAGAA 900
Qy 901 CAATTCCTTAACGATCCATATTTAATAGATCACTACTCTGAAACCGCTACAAATAACAAT 960
Db 901 CAATTCCTTAACGATCCATATTTAATAGATCACTACTCTGAAACCGCTACAAATAACAAT 960
Qy 961 CTTTCGATTACATCAAAAGGAAATATATCGTTGGATGGAGATTAGAACATTTCAAAAAA 1020
Db 961 CTTTCGATTACATCAAAAGGAAATATATCGTTGGATGGAGATTAGAACATTTCAAAAAA 1020
Qy 1021 ACCGATAATCTAGTCTATGTGATCTCCGTTTCGTTTATTTGTTGGGTAATGTGCA 1080
Db 1021 ACCGATAATCTAGTCTATGTGATCTCCGTTTCGTTTATTTGTTGGGTAATGTGCA 1080
Qy 1081 TTTTCTAAAGATGGCTAAATAAGTAGTGTTCGATGAAGAATTTAATCATTTGGGG 1140
Db 1081 TTTTCTAAAGATGGCTAAATAAGTAGTGTTCGATGAAGAATTTAATCATTTGGGG 1140
Qy 1141 GCGAAGATGTAGAAATTTGGTTACAGATTTATTTGCCAAAGGCTGTTTTTCAGAGTAAT 1200
Db 1141 GCGAAGATGTAGAAATTTGGTTACAGATTTATTTGCCAAAGGCTGTTTTTCAGAGTAAT 1200
Qy 1201 GACGGCGAATGGCCATCCATCAGAACACCCTGGTTAAGAAATGAACAGACGCGAA 1260
Db 1201 GACGGCGAATGGCCATCCATCAGAACACCCTGGTTAAGAAATGAACAGACGCGAA 1260
Qy 1261 GCTGTTAAAGTATTTACGCTTTAAATTTGTAAGAAAAAGGTACCTTACATCTATAGAAG 1320
Db 1261 GCTGTTAAAGTATTTACGCTTTAAATTTGTAAGAAAAAGGTACCTTACATCTATAGAAG 1320
Qy 1321 CTTTACCATAAGAGATTCACATATTCATAGAATACCTTTAGTTTCTATTATATCCCC 1380
Db 1321 CTTTACCATAAGAGATTCACATATTCATAGAATACCTTTAGTTTCTATTATATCCCC 1380
Qy 1381 GCTTATACCTGCAAAATATTTCAAGATGTGTAGATGCTCTTTAATCAAACTGTT 1440
Db 1381 GCTTATACCTGCAAAATATTTCAAGATGTGTAGATGCTCTTTAATCAAACTGTT 1440
Qy 1441 GTCGATCTCGAGGTTTGTATTGTTAAGCATGTTCAACAGATAATACCTTAGAAGTGATC 1500
Db 1441 GTCGATCTCGAGGTTTGTATTGTTAAGCATGTTCAACAGATAATACCTTAGAAGTGATC 1500

Qy 1501 AATAAGCTTTATGGTAATAATCCTAGGGTACGCATCATGTCTAAACCAATGCGGAATA 1560
Db 1501 AATAAGCTTTATGGTAATAATCCTAGGGTACGCATCATGTCTAAACCAATGCGGAATA 1560
Qy 1561 GCTCAGCATCAAAATCGACCGCTTTCTTTGCTAAAGGTTATTACATTGGCGAGTTAGAT 1620
Db 1561 GCTCAGCATCAAAATCGACCGCTTTCTTTGCTAAAGGTTATTACATTGGCGAGTTAGAT 1620
Qy 1621 TCAGATGATTATCTTCAGCTGATGAGTGAACCTGTTGTTTAAAGAAATTTTAAAGAT 1680
Db 1621 TCAGATGATTATCTTCAGCTGATGAGTGAACCTGTTGTTTAAAGAAATTTTAAAGAT 1680
Qy 1681 AAAACGCTAGCTTGTGTTTATACCACTAATAGAAACGTCATCCGGATGCTAGCTTAATC 1740
Db 1681 AAAACGCTAGCTTGTGTTTATACCACTAATAGAAACGTCATCCGGATGCTAGCTTAATC 1740
Qy 1741 GCTAATGTTTACAATTTGGCCAGAAATTTTCACGAGAAAACTCACACGGCTATGATTGCT 1800
Db 1741 GCTAATGTTTACAATTTGGCCAGAAATTTTCACGAGAAAACTCACACGGCTATGATTGCT 1800
Qy 1801 CACCAATTTAGAATGTTTACGATTAGAGCTTGGCAATTTAAACGGATGATTAAACGAAAT 1860
Db 1801 CACCAATTTAGAATGTTTACGATTAGAGCTTGGCAATTTAAACGGATGATTAAACGAAAT 1860
Qy 1861 ATTGAAACCGCTGGATTTATGACATGTTCCCTTAAACCTCAGTGAAGTTGGAATTTAAA 1920
Db 1861 ATTGAAACCGCTGGATTTATGACATGTTCCCTTAAACCTCAGTGAAGTTGGAATTTAAA 1920
Qy 1921 CATCTTAATAAATCTGCTATACCGGCTATTACATGGTGATACACATCCATTAAAGAAA 1980
Db 1921 CATCTTAATAAATCTGCTATACCGGCTATTACATGGTGATACACATCCATTAAAGAAA 1980
Qy 1981 CTGGCATTTCAAAAGAAAAACCAATTTTGTGTAGTCAATCAGTCAATTAATAGACAGGC 2040
Db 1981 CTGGCATTTCAAAAGAAAAACCAATTTTGTGTAGTCAATCAGTCAATTAATAGACAGGC 2040
Qy 2041 ATCAATTTATTAATTTATGACAAATTTGATGATTTAGATGAAAGTATATCTTC 2100
Db 2041 ATCAATTTATTAATTTATGACAAATTTGATGATTTAGATGAAAGTATATCTTC 2100
Qy 2101 AATAAAACCGCTGAATATCAAGAAAGAAATGGATATGTTTAAAGATCTTAAACCTCAAT 2160
Db 2101 AATAAAACCGCTGAATATCAAGAAAGAAATGGATATGTTTAAAGATCTTAAACCTCAAT 2160
Qy 2161 AATAAAAGTCCAAATTTACACCGAGCATCAAAAGAAATTTGGCTTTCTATCATAGCAC 2220
Db 2161 AATAAAAGTCCAAATTTACACCGAGCATCAAAAGAAATTTGGCTTTCTATCATAGCAC 2220
Qy 2221 AAAAACTTAAACATATTTATTAATAATAATAATAATAATAATAATAATAATAATAATA 2280
Db 2221 AAAAACTTAAACATATTTATTAATAATAATAATAATAATAATAATAATAATAATAATA 2280
Qy 2281 GATAAGATCATCTTACACCGAGCATCAAAAGAAATTTGGCTTTCTATCATAGCAC 2340
Db 2281 GATAAGATCATCTTACACCGAGCATCAAAAGAAATTTGGCTTTCTATCATAGCAC 2340
Qy 2341 CAAGTGAATTTTACTTAAATAATGACATCTCATATTTACACGAGTAATAGACTAATAAA 2400
Db 2341 CAAGTGAATTTTACTTAAATAATGACATCTCATATTTACACGAGTAATAGACTAATAAA 2400
Qy 2401 ACTGAGGCACATTTAAGTAATTAATAATAATAATAATAATAATAATAATAATAATAATA 2460
Db 2401 ACTGAGGCACATTTAAGTAATTAATAATAATAATAATAATAATAATAATAATAATAATA 2460
Qy 2461 ATCATTTTGTGATATCATGACAGCCTATTGTTTAAATGACAGCTATGCTTTATATGAA 2520
Db 2461 ATCATTTTGTGATATCATGACAGCCTATTGTTTAAATGACAGCTATGCTTTATATGAA 2520
Qy 2521 AATATGATGTCGGCATGAATTTCTCAGCATTAACACATGATTGGATCGAGAAAAATCAAT 2580
Db 2521 AATATGATGTCGGCATGAATTTCTCAGCATTAACACATGATTGGATCGAGAAAAATCAAT 2580

Qy	2581	GCGCATCCACCATTTAAANAGCTGATTAANACCTATTTTAATGACAAATGACTTAAGAAGT	2640
Db	2581	GCGCATCCACCATTTAAANAGCTGATTAANACCTATTTTAATGACAAATGACTTAAGAAGT	2640
Qy	2641	ATGAATGTCAAAAGGGGCATCACAAAGGTATGTTTATGAAGTATGCGCTACGCGCATGAGCTT	2700
Db	2641	ATGAATGTCAAAAGGGGCATCACAAAGGTATGTTTATGAAGTATGCGCTACGCGCATGAGCTT	2700
Qy	2701	CTGACGATTTATTAAGAAGTCATCATACATCCTGCCAATCAATTGTATAGTGTGCCAAGATAT	2760
Db	2701	CTGACGATTTATTAAGAAGTCATCATACATCCTGCCAATCAATTGTATAGTGTGCCAAGATAT	2760
Qy	2761	AACACTGAGGATATTTGGTTCCAATTTGCACCTTTTAATCTTTAGAAAAAGAAAACCGGGCAT	2820
Db	2761	AACACTGAGGATATTTGGTTCCAATTTGCACCTTTTAATCTTTAGAAAAAGAAAACCGGGCAT	2820
Qy	2821	GTATTTTAATTAACAATCGACCCCTGACTTATATGCTCTGGGAACGAAATATACAAATGGACA	2880
Db	2821	GTATTTTAATTAACAATCGACCCCTGACTTATATGCTCTGGGAACGAAATATACAAATGGACA	2880
Qy	2881	AATGAACAATATCAAAAGTCAAAAAAGGCGAAAAATATCCCGCTTAACAAGTTCATTATT	2940
Db	2881	AATGAACAATATCAAAAGTCAAAAAAGGCGAAAAATATCCCGCTTAACAAGTTCATTATT	2940
Qy	2941	AATAGTATACCGTATATAACAATTTGCAATTTTATTAATAA	2979
Db	2941	AATAGTATACCGTATATAACAATTTGCAATTTTATTAATAA	2979

RESULT 2

```

US-09-842-484A-3
; Sequence 3, Application US/09842484A
; GENERAL INFORMATION:
; APPLICANT: DEANGELIS, PAUL L.
; TITLE OF INVENTION: CHONDROITIN SYNTHASE GENE AND METHODS OF MAKING AND USING SAME
; FILE REFERENCE: 4605.003
; CURRENT APPLICATION NUMBER: US/09/842,484A
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2979
; TYPE: DNA
; ORGANISM: Pasteurella multocida
US-09-842-484A-3

```

[illegible]

Db	301		CTATCCGAATCAGAAAAACAGTTTAAATAAATAAATGAAATCTATCCTCGGAAAAA	360
Qy	361		TCGGAGAACGAGAAATCAGAAAGGTGGAACCTAGTAGTACCCAAAGATTTTCCCTAAAGATCTT	420
Db	361		TCGGAGAACGAGAAATCAGAAAGGTGGAACCTAGTAGTACCCAAAGATTTTCCCTAAAGATCTT	420
Qy	421		GTTCTTGCTCCATGTGCCAGATCATGTTTAATGATTTTACATGATGTCACAAAATCGAAAAAA	480
Db	421		GTTCTTGCTCCATGTGCCAGATCATGTTTAATGATTTTACATGATGTCACAAAATCGAAAAAA	480
Qy	481		AGCTTAGGTATAAAGCCCTGTAATAAAGAAATATCGGTCTTTCTATTATTATTTCCTACATTT	540
Db	481		AGCTTAGGTATAAAGCCCTGTAATAAAGAAATATCGGTCTTTCTATTATTATTTCCTACATTT	540
Qy	541		AATCGTAGCCGCTATTTTAGATATAACGTTTAGCCTCTTTGGTCAATCAGAAAAACAACCTAC	600
Db	541		AATCGTAGCCGCTATTTTAGATATAACGTTTAGCCTCTTTGGTCAATCAGAAAAACAACCTAC	600
Qy	601		CCATTTGAAGTCGTTGTTGCCAGATGATGGTAGTAAGGAAAACTTACTTACCATTGTGCAA	660
Db	601		CCATTTGAAGTCGTTGTTGCCAGATGATGGTAGTAAGGAAAACTTACTTACCATTGTGCAA	660
Qy	661		AAATACGAACAAAAAATTTGACATAAAGTATGTAAGACAAAAAGATTATGGGATATCAATTG	720
Db	661		AAATACGAACAAAAAATTTGACATAAAGTATGTAAGACAAAAAGATTATGGGATATCAATTG	720
Qy	721		TGTGCAGTCAAGAACTTAGTGGTTACGTACAGCAAAAGTATGATTTGTCTCGATCTCTAGAC	780
Db	721		TGTGCAGTCAAGAACTTAGTGGTTACGTACAGCAAAAGTATGATTTGTCTCGATCTCTAGAC	780
Qy	781		TGCGATATGGCACCACAAATATATGGGTTCATCTTATCTTACAGAACTATTAGAAGAC	840
Db	781		TGCGATATGGCACCACAAATATATGGGTTCATCTTATCTTACAGAACTATTAGAAGAC	840
Qy	841		AATGATATTGTTTAAATTTGGACCTAGAAAAATATGTGGATACATCAATAATTACCGCAGAA	900
Db	841		AATGATATTGTTTAAATTTGGACCTAGAAAAATATGTGGATACATCAATAATTACCGCAGAA	900
Qy	901		CAATTCCTTAACGATCCATATTTAATAGATCACTACCTGGAACCGCTACAATACAAAT	960
Db	901		CAATTCCTTAACGATCCATATTTAATAGATCACTACCTGGAACCGCTACAATACAAAT	960
Qy	961		CCTTCGATTACATCAAAAGGAAATATATCGTTGGATTTGGAGATTAGAATTTCAAAAAA	1020
Db	961		CCTTCGATTACATCAAAAGGAAATATATCGTTGGATTTGGAGATTAGAATTTCAAAAAA	1020
Qy	1021		ACCGATAATCTACGTCATATGATTTCTCGTTTTCGTTATTTTGTTCGGGGTAATGTGCA	1080
Db	1021		ACCGATAATCTACGTCATATGATTTCTCGTTTTCGTTATTTTGTTCGGGGTAATGTGCA	1080
Qy	1081		TTTTTCTAAGAAATGGCTAAATAAAGTACGTTGGTTCGATGAAGAATTAATCATTTGGGG	1140
Db	1081		TTTTTCTAAGAAATGGCTAAATAAAGTACGTTGGTTCGATGAAGAATTAATCATTTGGGG	1140
Qy	1141		GGCGAAGATGTAGAAATTTGGTTTACAGATTATTTGGCAAGGCTGTTTTTCAGAGTAAAT	1200
Db	1141		GGCGAAGATGTAGAAATTTGGTTTACAGATTATTTGGCAAGGCTGTTTTTCAGAGTAAAT	1200
Qy	1201		GACGCGGAATGGCCATCCATCAAGAACCCCTGGTGAAGAAAAATGAACAGAACCGCAA	1260
Db	1201		GACGCGGAATGGCCATCCATCAAGAACCCCTGGTGAAGAAAAATGAACAGAACCGCAA	1260
Qy	1261		GCTGCTAAGATGATTAACGCTTAAAAATGTGAAGAAAAAGGTACCTTACATCTATAGAAAG	1320
Db	1261		GCTGCTAAGATGATTAACGCTTAAAAATGTGAAGAAAAAGGTACCTTACATCTATAGAAAG	1320
Qy	1321		CTTTTACCATAAGAAAGTTTACATATTCATAGAATACCTTTTAGTTCTTATTTATATCCTCC	1380
Db	1321		CTTTTACCATAAGAAAGTTTACATATTCATAGAATACCTTTTAGTTCTTATTTATATCCTCC	1380
Qy	1381		GCTTATAACTGTGCAAAATTTATTTCAAGATGTGTAGATAGTGCCTTTAATCAAACTGTT	1440

Db 1381 GCTTATACTGTGCAAAATATATATCAAGATGCTAGTAGTGCTCTTAATCAAACTGTT 1440
QY 1441 GTCGATCTCGAGGTTGTATTTGAACGATGTTTCAACAGATAATACCTTAGAAGTGATC 1500
Db 1441 GTCGATCTCGAGGTTGTATTTGAACGATGTTTCAACAGATAATACCTTAGAAGTGATC 1500
QY 1501 AATAAGCTTTATGTAATAATCCTAGGTPACGCATCATCTCTAAACCAATGGCGGAATA 1560
Db 1501 AATAAGCTTTATGTAATAATCCTAGGTPACGCATCATCTCTAAACCAATGGCGGAATA 1560
QY 1561 GCGTCAGCATCAAAATGACGCCGTTCTTTTGTAAAGGTTATTAATGCGGAGTTAGAT 1620
Db 1561 GCGTCAGCATCAAAATGACGCCGTTCTTTTGTAAAGGTTATTAATGCGGAGTTAGAT 1620
QY 1621 TCAGATGATATCTTCAGGCTGATGAGTTGAACCTGTGTTAAAGAAATTTTAAAGAT 1680
Db 1621 TCAGATGATATCTTCAGGCTGATGAGTTGAACCTGTGTTAAAGAAATTTTAAAGAT 1680
QY 1681 AAAACGCTAGCTGTGTTTATACCACTAATAGAAACGTCCTCAATCCGGATGGTAGCTTAATC 1740
Db 1681 AAAACGCTAGCTGTGTTTATACCACTAATAGAAACGTCCTCAATCCGGATGGTAGCTTAATC 1740
QY 1741 GCTAATGTTTACAAATGGCGAGAAATTTTCACGAGAAAACTCACACGGCTATGATGCT 1800
Db 1741 GCTAATGTTTACAAATGGCGAGAAATTTTCACGAGAAAACTCACACGGCTATGATGCT 1800
QY 1801 CACCAATTTAGAAATGTTTACGATTTAGAGCTTGCCATTTAACGGATGGATTTAACGAAAT 1860
Db 1801 CACCAATTTAGAAATGTTTACGATTTAGAGCTTGCCATTTAACGGATGGATTTAACGAAAT 1860
QY 1861 ATTGAAACGCCGTGGATTTAGCATGTTCCCTTAACTCAGTGAAGTTGGAATTTAAA 1920
Db 1861 ATTGAAACGCCGTGGATTTAGCATGTTCCCTTAACTCAGTGAAGTTGGAATTTAAA 1920
QY 1921 CATCTTAATAAATCTGCTATATACCGGTATATACATGCTGATACACATCCATTAAAGAA 1980
Db 1921 CATCTTAATAAATCTGCTATATACCGGTATATACATGCTGATACACATCCATTAAAGAA 1980
QY 1981 CTGCGCATTCAAAAGAAAAACCAATTTTGTGTAGTCAATCAGTCATTTAAATAGACAAGC 2040
Db 1981 CTGCGCATTCAAAAGAAAAACCAATTTTGTGTAGTCAATCAGTCATTTAAATAGACAAGC 2040
QY 2041 ATCAATTTATTAATTAATGACAAATTTGATGATTTAGATGAAAGTAAAGTATATCTTC 2100
Db 2041 ATCAATTTATTAATTAATGACAAATTTGATGATTTAGATGAAAGTAAAGTATATCTTC 2100
QY 2101 AATAAAACCGCTGAATATCAAGAGAAATGGATATGTTTAAAGATCTTAAACTCATTCAA 2160
Db 2101 AATAAAACCGCTGAATATCAAGAGAAATGGATATGTTTAAAGATCTTAAACTCATTCAA 2160
QY 2161 AATAAGATGCAAAATCGCAGTCAGTATTTCTATCCCAATACATTTAAACGGCTTAGTG 2220
Db 2161 AATAAGATGCAAAATCGCAGTCAGTATTTCTATCCCAATACATTTAAACGGCTTAGTG 2220
QY 2221 AAAAACTAAACATATATTTGAATATATAAATAATATTCGTTATTTATCTACATGTT 2280
Db 2221 AAAAACTAAACATATATTTGAATATATAAATAATATTTCTATCCCAATACATTTAAACGGCTTAGTG 2280
QY 2281 GATAAGATCATCTTACACAGACATCAAAAAGAAATATTTGGCTTTCTATCATAAGCAC 2340
Db 2281 GATAAGATCATCTTACACAGACATCAAAAAGAAATATTTGGCTTTCTATCATAAGCAC 2340
QY 2341 CAAGTGAATATTTACTAAATTAATGACATCTCATATTTACAGAGTAAATAGACTAATAAA 2400
Db 2341 CAAGTGAATATTTACTAAATTAATGACATCTCATATTTACAGAGTAAATAGACTAATAAA 2400
QY 2401 ACTGAGGACATTTAAGTAATTAATAAATTAAGTCAAGTTAAATCTAAATTTGTAATAC 2460
Db 2401 ACTGAGGACATTTAAGTAATTAATAAATTAAGTCAAGTTAAATCTAAATTTGTAATAC 2460
QY 2461 ATCAATTTTGTATATCATGACGCCATTCGTTAAAAATGACAGCTATGCTTTATATGAAA 2520
Db 2461 ATCAATTTTGTATATCATGACGCCATTCGTTAAAAATGACAGCTATGCTTTATATGAAA 2520

QY 2521 AAATATGATGTGGCATGAATTTCTCAGCATTAAACACATGATTTGGATCGAGAAAAATCAAT 2580
Db 2521 AAATATGATGTGGCATGAATTTCTCAGCATTAAACACATGATTTGGATCGAGAAAAATCAAT 2580
QY 2581 GGGCATCCACCATTTAAAAGCTGATTAACACCTATTTTAATGACATGACTTTAAGAAGT 2640
Db 2581 GGGCATCCACCATTTAAAAGCTGATTAACACCTATTTTAATGACATGACTTTAAGAAGT 2640
QY 2641 ATGAATGTAAAGGGGCATCACAAAGTATGTTTATGAAGTATCGCTACCGCATGAGCTT 2700
Db 2641 ATGAATGTAAAGGGGCATCACAAAGTATGTTTATGAAGTATCGCTACCGCATGAGCTT 2700
QY 2701 CTGACGATTTATAAGAAAGTCATCACATCTGCCAATCAATTTGATGTTGCCAGATAT 2760
Db 2701 CTGACGATTTATAAGAAAGTCATCACATCTGCCAATCAATTTGATGTTGCCAGATAT 2760
QY 2761 AACACTGAGGATATTTGGTTCCAAATTTGCACTTTTAACTTTAGAAAAAGAAACCGGCAT 2820
Db 2761 AACACTGAGGATATTTGGTTCCAAATTTGCACTTTTAACTTTAGAAAAAGAAACCGGCAT 2820
QY 2821 GTATTTAATAAAACATCGACCCCTGACTTATATGCTTTGGGAACGAAATTTACATGGACA 2880
Db 2821 GTATTTAATAAAACATCGACCCCTGACTTATATGCTTTGGGAACGAAATTTACATGGACA 2880
QY 2881 AATGAACAAATTTCAAAGTGCAGAAAAAGGCGAAATATCCCGTTAACAAAGTTTCATTAT 2940
Db 2881 AATGAACAAATTTCAAAGTGCAGAAAAAGGCGAAATATCCCGTTAACAAAGTTTCATTAT 2940
QY 2941 AATAGTATAAGCTTATAAAACATTTGCACTTTTATTTAAA 2979
Db 2941 AATAGTATAAGCTTATAAAACATTTGCACTTTTATTTAAA 2979

RESULT 3
US-09-469-200D-8
; Sequence 8, Application US/09469200D
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H.
; APPLICANT: DeAngelis, Paul
; APPLICANT: Kumari, Kshama
; TITLE OF INVENTION: Hyaluronan Synthase Gene and Uses Thereof
; FILE REFERENCE: 3554.011
; CURRENT APPLICATION NUMBER: US/09/469,200D
; PRIOR FILING DATE: 1999-12-21
; PRIOR FILING DATE: 1998-10-26
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 2937
; TYPE: DNA
; ORGANISM: Pastuerella Multocida
US-09-469-200D-8

Query Match 76.7%; Score 2285; DB 5; Length 2937;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 2541; Conservative 0; Mismatches 375; Indels 21; Gaps 1;

QY 43 ATTTTAAAGGAAGAAATGAATATATCACAAGCAATAAAGCATATTAACAGCAAT 102
Db 1 ATTTTAAAGGACACAAAATGAATATATCACAAGCAATAAAGCATATTAACAGCAAT 60
QY 103 GACTATGATTTAGCACTCAAAATTTTGAAGTCTGCTGAACCTACCGGGGAAAAATC 162
Db 61 GACTATCAATTTAGCACTCAAAATTTTGAAGTCTGCTGAACCTACCGGGGAAAAATC 120
QY 163 GTTGAATTTCAAAATTTATCAAAATGTAAGAAAAAATCTCTGACCAATTC----- 209
Db 121 GTTGAATTTCAAAATTTATCAAAATGTAAGAAAAAATCTCTGACCAATTCCTCTGTTAATTC 180

Qy	210	-----TTATGTAGTGAAGATATAAAAAACAGTGTTCGGATAGCTCAATTAGATATC	261
Db	181	GCACATCTTCTCTTAATAAAGAAGAAAGTCAATGTTCGGATATGTCGGTTAGATATT	240
Qy	262	GCAACACAGCTCTTACTTTCACAGTAAAAAATAAAGTCTATCCGAATCAGAAAAAAC	321
Db	241	GCAACACACTGTACTTTCACAGTAAAAAATAAAGTCTATCCGAATCAGAAAAAAC	300
Qy	322	AGTTTAAAAAATAAATGGAAATCTACTCGGAAAAATCGGAGAACCGAGAAATCAGA	381
Db	301	ACGTTAAAAAATAAATGGAAATCTACTCGGAAAAATCGGAGAAATCTGGAATCGGAGGTAGA	360
Qy	382	AAGGTGGAATAGTACCCAAAGATTTTCTCAAGATCTTGCTTGTCTGCTCAATGCGAGAT	441
Db	361	CGCGTCGCCCTTGTACCAAAAGATTTTCCCAAGATCTGGTTTTAGCCGCTTTACCTGAT	420
Qy	442	CATGTTAATCATTTTACATGGGTACAAAAATCGAAAAAAGCTTAGGTATAAAGCCCTGA	501
Db	421	CATGTTAATCATTTTACATGGGTACAAAAAGGAAAGAAAGACTTGGCATATAAACCTTGAA	480
Qy	502	AATAAGAAATACGGTCTTCTTATTTATTTCTACATTTAATCGTAGCGGTATTTTTAGAT	561
Db	481	CATCAACATGTGCTTCTTCTATTTATTTCAACATTTCAATCGACACGAATTTTATCG	540
Qy	562	ATAAGTGTAGCCTTGTGTCATCAGAAAAACAATACCCAAATTTGAAGTCTGTTGTGCGA	621
Db	541	ATTACATTTAGCCTTGTGTTAAACCAAAAAACACATATACCCGCTTTGAAGTTATCGTGACA	600
Qy	622	GATGATGGTAGTAAGAAAAACTTACTTACCATTGTGCAAAAATACGAACAAAACTTGAC	681
Db	601	GATGATGGTAGTCAGGAAGATCTATCACCGATCATTCGCCAAATATGAAATTAATTTGGAT	660
Qy	682	ATAAAGTATCTAAGACAAAAAGATTATGATATCAATTTGTGTCAGTCAGAAAACTTAGGT	741
Db	661	ATTGCTGTAGCTCAGACAAAAAGATACGGTTTTCAAGCCAGTCCGCTCGGAATATGGGA	720
Qy	742	TTAGTACAGCAAAAGTATGATTTTGTCTCGATTCTAGACTCGCATATGGCACCAACAAC	801
Db	721	TTAGCCTTAGCAAAATATGACTTTATTTGGCTTACTCGACTGTGATATGGCGCAAAATCCA	780
Qy	802	TTATGGGTTCATTTCTTACAGAACTATTAGAAGACAAATGATATTGTTTTAATTTGGA	861
Db	781	TTATGGGTTCATTTCTTATGTTGAGAGCTATTAGAAGATGATGATTTAAACAATCATTTGGT	840
Qy	862	CCTAGAAAAATATGTGGTACTCATATAATTACCCGAGAACAAATTCCTTTAAGCATCCATAT	921
Db	841	CCAAGAAATACATCGATACACAACATATTGACCCCAAGACTTCTTTAATAAGCCGAGT	900
Qy	922	TTAATAGAACTACTACTGAAACCGGTCAAAATAACAATCTCTGATTTACATCAAAAAGGA	981
Db	901	TTGCTTGAATCATTTACAGAAGTGA AAAACCAATAATAGTTGTGCGCAAAAAGGGGAAGGA	960
Qy	982	AATATATCGTTGATTTGGAGATTAGAACATTTTCAAAAAACCGGATAATCTACGCTTATGT	1041
Db	961	ACAGTTTCTCTGATTTGGCGCTTTAGAACAAATTCGAAAAACAGAAATATCTCGCTTATCC	1020
Qy	1042	GATTCCTCGTTTCGTTATTTTGTGCGGTAAATGTTGCAATTTTCTTAAAGATGGCTAAAT	1101
Db	1021	GATTCGCTTTCCGTTTTTTTTTGGCGCGGTAAATGTTGCTTTGAAATAATGGCTAAAT	1080
Qy	1102	AAAGTAGGTTGGTTCGATGAAGAAATTAATCATTTGGGGGGCGGAAGATGTAGAATTTGGT	1161
Db	1081	AAATCCGGTTCTTGTAGAGGAATTTAATCACTGGGTGGAGAGAGATGTGAATTTGGA	1140
Qy	1162	TACAGATATTTGCCAAAGCTTTTTTTTCAGAGTAATTTGACGGCGGAATGGCCATCCAT	1221
Db	1141	TATCGCTTATTCGGTTACGGTAGTTTCTTTTAAACATAATGTTGCTTTGAAATAATGGCTAAAT	1200
Qy	1222	CAAGAACCACTCGTAAGAAAAATGAACAGAACCGGAGCTGGTTAAAAGTATTACGCTT	1281
Db	1201	CAAGGCCACCGGTAAAGAAAAATGAACCGCATCGTGAAGCGGGA AAAAATATTACGGCT	1260
Qy	1282	AAAAATTGAAAAAAGGTACTCTTACATCTATAGAAAGCTTTTACCNAATAGAGATTCA	1341

[illegible]

Qy 1282 AAAATTGTGAAAGAAAGGTACCTTACATCTATAGAAAGCTTTTACCAATAGAGATTCA 1341
Db 1261 GATATTATGAGAGAAAGGTCCCTTATATCTATAGAAACCTTTTACCAATAGAGATTTCG 1320
Qy 1342 CATATTATAGAAATACCTTTTATAGTTTCTATTATATATCCCGCTTATAACTGTGCAAAATAT 1401
Db 1321 CATATCAATAGAGTACCTTTTATAGTTTCAATTTTATATCCAGCTTATAACTGTGCAAACTAT 1380
Qy 1402 ATTCAAAGATCTGTAGATAGCTTCTTAATCAAACTGTGTGCGATCTCGAGGTTTGTATT 1461
Db 1381 ATTCAAGCTTGGTGTAGATAGTGCCTGAATCAGACTGTGTGATCTCGAGGTTTGTATT 1440
Qy 1462 TGTAAACGATGGTTCACACAGATAATACCTTTAGAAGTGATCAATAAGCTTTATGGTAATAAT 1521
Db 1441 TGTAAACGATGGTTCACACAGATAATACCTTTAGAAGTGATCAATAAGCTTTATGGTAATAAT 1500
Qy 1522 CCTAGGTTACGCATCATGTCTAAACCAATGGCGGAATAGCCTCAGCATCAAAATGCAGCC 1581
Db 1501 CCTAGGTTACGCATCATGTCTAAACCAATGGCGGAATAGCCTCAGCATCAAAATGCAGCC 1560
Qy 1582 GTTTCCTTTTGTAAAGGTTATTACATTTGGGCGAGTTAGATTTCAGATGATTATCTTGAGCCT 1641
Db 1561 GTTTCCTTTTGTAAAGGTTATTACATTTGGGCGAGTTAGATTTCAGATGATTATCTTGAGCCT 1620
Qy 1642 GATCAGCTTGAACCTGTGTTTAAAGAAATTTTAAAGATAAAACGCTAGCTTTGTGTTTAT 1701
Db 1621 GATCAGCTTGAACCTGTGTTTAAAGAAATTTTAAAGATAAAACGCTAGCTTTGTGTTTAT 1680
Qy 1702 ACCACTAATAGAAACGTCATATCCGGATGGTATGCTTAATFCGCTAATGGTTACAAATGGCCA 1761
Db 1681 ACCACTAATAGAAACGTCATATCCGGATGGTATGCTTAATFCGCTAATGGTTACAAATGGCCA 1740
Qy 1762 GAATTTTCAGAGAAACCTCACACGCTATGATGCTCACCATTTTAGATGTTTACG 1821
Db 1741 GAATTTTCAGAGAAACCTCACACGCTATGATGCTCACCATTTTAGATGTTTACG 1800
Qy 1822 ATTAGAGCTTGGCATTTTAAACGGATGGATTTAACGAAAATATTTGAAACGCCGTGATTAT 1881
Db 1801 ATTAGAGCTTGGCATTTTAAACGGATGGATTTAACGAAAATATTTGAAACGCCGTGATTAT 1860
Qy 1882 GACATGTTCTTAACTCAGTGAAGTTGGAAAAATTTAAACATCTTAATAAATCTGCTAT 1941
Db 1861 GACATGTTCTTAACTCAGTGAAGTTGGAAAAATTTAAACATCTTAATAAATCTGCTAT 1920
Qy 1942 AACCGCTATTACATGTTGTAACACATCCATTAAGAACTCGGCATTCACAAAGAAAAAC 2001
Db 1921 AACCGCTATTACATGTTGTAACACATCCATTAAGAACTCGGCATTCACAAAGAAAAAC 1980
Qy 2002 CATTTTGTGTAGTCAATCAGTCAATTAATAGAACGTCATCAATTTATTAATATATGAC 2061
Db 1981 CATTTTGTGTAGTCAATCAGTCAATTAATAGAACGTCATCAATTTATTAATATATGAC 2040
Qy 2062 AAATTTGATGATTTAGATGAAGTAGAAGTATATCTTCATATAAACCGCTGAATATCAA 2121
Db 2041 GAATTTGATGATTTAGATGAAGTAGAAGTATATTTTCAATAAACCGCTGAATATCAA 2100
Qy 2122 GAAGAAATGGATGTTTAAAGATCTTAAACTCATTTCAAAATAAAGATGCCAAATCGCA 2181
Db 2101 GAAGAGATTGATCTTAAAGATATTAATATCATCCAGATTAAGATGCCAAATCGCA 2160
Qy 2182 GTCAGTATTTTCTATCCCAATACATTAACCGCTTAGTGAAGAACTAAACAAATATTAT 2241
Db 2161 GTCAGTATTTTCTATCCCAATACATTAACCGCTTAGTGAAGAACTAAACAAATATTAT 2220
Qy 2242 GAATATATAAATAATATTCGTTATTATTTCTAGATGTTGATAGAAATCATCTTACACCA 2301
Db 2221 GAATATATAAATAATATTCGTTATTATTTCTAGATGTTGATAGAAATCATCTTACACCA 2280
Qy 2302 GACATCAAAAAGAAATATTCGCTTTCTATCATAGCACCAGTGAATATTTTACTAAAT 2361
Db 2281 GATATCAAAAAGAAATATTCGCTTTCTATCATAGCACCAGTGAATATTTTACTAAAT 2340

Qy 2362 AATGACATCTCATATTACACGAGTAATAGACTAATAAAAACTGAGCACATTTAAGTAAT 2421
Db 2341 AATGATATCTCATATTACACGAGTAATAGACTAATAAAAACTGAGCGCATTTAAGTAAT 2400
Qy 2422 ATTAATAAATTAAGTCAGTTAAATCTAAATGTGAATACATCATTTTGTATATCATGAC 2481
Db 2401 ATTAATAAATTAAGTCAGTTAAATCTAAATGTGAATACATCATTTTGTATATCATGAC 2460
Qy 2482 AGCCTATTTCGTTAAATAATGACAGCTATGCTTATATGAAAAATATCATCTCGCATGAAT 2541
Db 2461 AGCCTATTTCGTTAAATAATGACAGCTATGCTTATATGAAAAATATCATCTCGCATGAAT 2520
Qy 2542 TTCTCAGCATTAACACATGATTGGATCGAGAAAACTAATCGCATCCACCATTTAAAAAG 2601
Db 2521 TTCTCAGCATTAACACATGATTGGATCGAGAAAACTAATCGCATCCACCATTTAAAAAG 2580
Qy 2602 CTGATTAAACCTTATTAAATGACAATGACTTAAAGAATGATGAATGTGAAAGGGGCATCA 2661
Db 2581 CTGATTAAACCTTATTAAATGACAATGACTTAAAGAATGATGAATGTGAAAGGGGCATCA 2640
Qy 2662 CAAGGTATGTTTATGAAGTATGCGTACCGCATGAGCTTCTGACGATTTTAAAGAAGTC 2721
Db 2641 CAAGGTATGTTTATGACGATGCGTACCGCATGAGCTTCTGACGATTTTAAAGAAGTC 2700
Qy 2722 ATCACATCTCGCATCAATTTGATAGTGTGCGAGAAATATAACACTGAGGATATTTGGTTC 2781
Db 2701 ATCACATCTCGCATCAATTTGATAGTGTGCGAGAAATATAACACTGAGGATATTTGGTTC 2760
Qy 2782 CAATTTGACATTTTAACTTAGAAAAAGAAACCGGCATGTTATTAAACACATCGACC 2841
Db 2761 CAATTTGACATTTTAACTTAGAAAAAGAAACCGGCATGTTATTAAACACATCGACC 2820
Qy 2842 CTGACTTATATGCTTGGGACGAAATTTACAAATGACAAATTCACAAATTCACAACTGCA 2901
Db 2821 CTGACTTATATGCTTGGGACGAAATTTACAAATGACAAATTCACAAATTCACAACTGCA 2880
Qy 2902 AAAAAAGCGAAATATATCCCGTTAAACAGTTTCATTATTAAATAGTATAACGCTATAA 2958
Db 2881 AAAAAAGCGAAATATATATACCTGTTAAACAGTTTCATTATTAAATAGTATAACGCTATAA 2937

RESULT 5
US-10-217-613-2
; Sequence 2, Application US/10217613
; GENERAL INFORMATION:
; APPLICANT: DEANGELIS, PAUL
; TITLE OF INVENTION: DNA ENCODING HYALURONAN SYNTHASE FROM PASTURELLA MULTOCIDA ANI
; FILE REFERENCE: 35541.081
; CURRENT APPLICATION NUMBER: US/10/217,613
; PRIOR FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/283,402
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2920
; TYPE: DNA
; ORGANISM: Pasteurella multocida
US-10-217-613-2

Query Match 76.2%; Score 2271.2; DB 6; Length 2920;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 2526; Conservative 0; Mismatches 373; Indels 21; Gaps 1;
Qy 61 ATGAATACATTTATCAACAGCAATAAAGCATATATACAGCAATGACTATGAATTAGCACTC 120
Db 1 ATGAATACATTTATCAACAGCAATAAAGCATATATACAGCAATGACTATGAATTAGCACTC 60

QY 121 AAATTTTGAAGTCTGCTGAAACCTACGGCGGAAATAATCGTTGAATTCCAAATATC 180
DB 61 AAATTTTGAAGTCTGCGGAAATCTATGACGGAAATTTGTTGAATTTCAATATACC 120
QY 181 AAATGTAAGAAATACTCTCGACCAATCTTATG-----TAAGT 219
DB 121 AAATGCAAGAAATACTCTCAGCACATCTCTGTTAATTCAGCACATCTTCTGTGAAT 180
QY 220 GAAGATAAAAAACAGTGTTCGATAGCTCATTTAGATATCGCAACACAGCTCTTACAT 279
DB 181 AAAGAAGAAATACTCAATGTTTCGATAGCTCGTTAGATATTTGCAACACAACTGTTACAT 240
QY 280 TCCACGTAATAAATAATTAATCTATCCGAATCAGAAAAAACAAGTTTAAAAATAAATGG 339
DB 241 TCCACGTAATAAATAATTAATCTCTGACTCGGAATAAACAAGTTTAAAAATAAATGG 300
QY 340 AAATCTATCACTGGGAAAAATCGGAGAACGCGAGAAATCAGAAAGGTGGAACATAGTACCC 399
DB 301 AAATCTCACTGAGAAGAAATCTGAAATGCGGAGGTGAAGAGCGCTCGCCCTGTACCA 360
QY 400 AAGATTTTCCCTAAGATCTTGTCTTGTCCATTCGCCAGATCATGTTAATGATTTTACA 459
DB 361 AAGATTTTCCCAAGATCTGTTTGTAGCGCTTTACCTGATCATGTTAATGATTTTACA 420
QY 460 TGGTACAAAATCGAAAAAAGCTTAGGTATAAGCCCTGTAATAAGAAATATCGGCTCT 519
DB 421 TGGTACAAAAGCGAAGAAAGACTTGGCATAAAACCTGAACATCAACATGTTGCTCT 480
QY 520 TCTATTATTCTCATTTAATCGTAGCCGTATTTTAGATATTAAGTATGATGTTAGCTGTTG 579
DB 481 TCTATTATCTTACAACATTCATCGACACCAATTTTATFCGATTTACATTAGCTGTTA 540
QY 580 GTCAATCAGAAAAACAATCCCATTTGAAGTCGTTGTCAGATGATGTTAGTAAAGAA 639
DB 541 GTAAACAAAACACATTAACCGTGTGAAGTTATCGTGACAGATGATGCTAGCAGAA 600
QY 640 AACTTACTTACCATTGTGCAAAAATAGCAACAAAACCTTGACATAAAGTATGTAAGCAA 699
DB 601 GATCTATACCGGATTCAGCCCAATGAAATAAATTTGGATATTCGCTACGTCAGCAAA 660
QY 700 AAGATTTATGATATCAATTTGTGCGAGTCAGAAACCTTAGGTTTACGTACAGCAAAAT 759
DB 661 AAGATTAACGGTTTCAAGCAGTGGCGCTCGGAATATGGAATTTACGCTTAGCAAAAT 720
QY 760 GATTTTGTCTCGATTTAGACTCGGATATGGCACCAACAATATGAGGTTTCATTTCTAT 819
DB 721 GACTTTATGCTTACTCGACTGATATGGCGCAATCCATTTATGGTTTCATTTCTAT 780
QY 820 CTTACAGAACTATTAGAAGCAATGATATGTTTAAATGGACCTAGAAATATGTCGAT 879
DB 781 GTTCAGAGCTATTAGAAGATGATGATTTAACCAATCATTTGGTCCAGAAATATCATCGAT 840
QY 880 ACTCATTAATACCGCAGAACATTCCTTAACGATCCATATTTAATAGAACTACTACCT 939
DB 841 ACACAACATATTGACCCAAAGACTCTTAAATAACCGGAGTTTGGCTGATCATTTACCA 900
QY 940 GAAACCGCTACAAATAACAATCTCTCGATTTACATCAAAAGGAAATATATCGTTGATGG 999
DB 901 GAAGTGAACCAATAATAGTTTGGCGCAAAAGGGAAGGAAACAGTTTCTCTGGATTGG 960
QY 1000 AGATTAGAACATTTCAAAAACCGAATNCTACGTCTATGTGATTCCTCGTTTTCGTTAT 1059
DB 961 CGCTTTAGAACAAATTCGAAAAAACAGAAATCTCCGCTTATCCGATTCGCCCTTCGTTT 1020
QY 1060 TTTGTTGGGTAATGTTGCAATTTCTAAAGATGGCTAAATAAGTAGGTTGGTTTCGAT 1119
DB 1021 TTTGGGGGGTAAATGTTGCTTTCGCTAAAAAATGGCTAAATAAATCCGCTTCTTTGAT 1080
QY 1120 GAAGAAATTAATCATTTGGGGGGCGAGATGTAAGATTTGGTTACAGATTTATTTGCCAAA 1179
DB 1081 GAGGAATTTAATCATCTGGGGTGGAGAAGATGGAATTTGGATATCGCTTATTCGGTTAC 1140
QY 1180 GGCTGTTTTTTCAGAGTAATTTGACGGCGGAATGGCCATCCATCAAGAACCCCTGGTAAA 1239

DB 1141 GGTTAGTTTCTTTAAAACTATTGATGCAATATGSCCTACCATCAAGAGCCACAGGTAAA 1200
QY 1240 GAAATTAACACGAAACGCGAAGCTGTAAAGATTATACGCTTAAATTTGTAAGAAAAAG 1299
DB 1201 GAAATTAACACGCGATCGTGAAGCGGGAATAATATTACGCTCATATTATGAGAGAAAG 1260
QY 1300 GTACCTTACATCTATAGAAGCTTTTACCAATAGAAGATTACATATTCATAGAATACCT 1359
DB 1261 GTCCCTTATATCTATAGAAACCTTTTACCAATAGAAGATTCGCATATCAATAGATACCT 1320
QY 1360 TTAGTTTCTATTATATCCCGCTTATACTGTGCAAAATTAATTTCAAGATGTGTAGAT 1419
DB 1321 TTAGTTTCAATTTATATCCAGCTTAACTGTGCAAACTATATTCAACGTTCCGTTAGAT 1380
QY 1420 AGTGCTTTAATCAAACTGTTGTCGATCTCGAGSTTGTATTTGTAACTAGATGTTCAACA 1479
DB 1381 AGTGCACTGAATCAGACTGTTGTTGATCTCGAGSTTGTATTTGTAACTAGATGTTCAACA 1440
QY 1480 GATAATACCTTAGAAGTGATCAATAAGCTTTTATGGTAATATCTAGGGTACGCATCATG 1539
DB 1441 GATAATACCTTAGAAGTGATCAATAAGCTTTTATGGTAATATCTAGGGTACGCATCATG 1500
QY 1540 TCTAAACCAATGGCGGAATAGCTCAGCATCAATCGACCGCTTCTTTTGTCTAAAGGT 1599
DB 1501 TCTAAACCAATGGCGGAATAGCTCAGCATCAATCGACCGCTTCTTTTGTCTAAAGGT 1560
QY 1600 TATTACATTTGGGAGTTAGATTTCAGATGATTATCTTTCAGGCTGATGAGTGAATCTGTT 1659
DB 1561 TATTACATTTGGGAGTTAGATTTCAGATGATTATCTTTCAGGCTGATGAGTGAATCTGTT 1620
QY 1660 TTTAAAGAAATTTTAAAGATAAACCTAGCTTGTGTTTATACCACATTAATAGAAGCTC 1719
DB 1621 TTTAAAGAAATTTTAAAGATAAACCTAGCTTGTGTTTATACCACATTAATAGAAGCTC 1680
QY 1720 AATCCGATGCTAGCTTAATCGCTTAATGTTTACAATTTGGCCAGAAATTTTCACAGAAAA 1779
DB 1681 AATCCGATGCTAGCTTAATCGCTTAATGTTTACAATTTGGCCAGAAATTTTCACAGAAAA 1740
QY 1780 CTCACAAACGCTATGATTGCTCACCATTTTGAAGATTTTACGATTTAGAGCTTGGCATTTA 1839
DB 1741 CTCACAAACGCTATGATTGCTCACCATTTTGAAGATTTTACGATTTAGAGCTTGGCATTTA 1800
QY 1840 ACGGATGATTTTACGAAATATTTGAAACCGCTGATTTGACATGTTCTCTTAAACCTC 1899
DB 1801 ACTGATGATTTTCAATGAAATAATTTGAAATTTCCGCTAGACTATGACATGTTCTCTCAAACTC 1860
QY 1900 AGTGAAGTTGGAAAAATTTAAACATCTTAAATAAAATCTGCTATAACCGCTATTACATGGT 1959
DB 1861 AGTGAAGTTGGAAAAATTTAAACATCTTAAATAAAATCTGCTATAACCGCTATTACATGGT 1920
QY 1960 GATAACACATCTCATTAAGAACTCGGCAATTCAAAAGAAAAACCAATTTTGTGTAGTCAAT 2019
DB 1921 GATAACACATCAATTAAGAACTTTGGCAATTCAAAAGAAAAACCAATTTTGTGTAGTCAAT 1980
QY 2020 CAGTCATTAATACAGAGGCAATCAATTTATTAATTAATGACAAATTTGATGATTTAGAT 2079
DB 1981 CAGTCATTAATACAGAGGCAATCAATTTATTAATTAATGACAAATTTGATGATTTAGAT 2040
QY 2080 GAAAGTAGAAAGTATATCTTCAATAAAACCGCTGAATATCAAGAGAAATGGATGATGTA 2139
DB 2041 GAAAGTAGAAAGTATATTTTCAATAAAACCGCTGAATATCAAGAGAGATTTGATATCTTA 2100
QY 2140 AAAATCTTAAACATTCATAAATAAAGATGCCAAATCGCAGTCAGATTTTCTATCC 2199
DB 2101 AAAATCTTAAACATTCATAAATAAAGATGCCAAATCGCAGTCAGATTTTCTATCC 2160
QY 2200 AATACATTTAAACGGCTTAGTGAAAAAATAAACAATAATTTATTAATAATAAATAATA 2259
DB 2161 AATACATTTAAACGGCTTAGTGAAAAAATAAACAATAATTTATTAATAATAAATAATA 2220
QY 2260 TTCGTTATTTTCTACATGTTGATAAGATCATCTTACACAGACATCAAAAAGAAATA 2319

Db 2221 TTCGTTATTGTTTACATGTTGATAAGAAATCATCTTACACAGATATCAAAAAAGAAATA 2280
Qy 2320 TTGGCTTTCTATCAAGCAACCAAGTGAATATTTTACTAAATAATGACATCTCATATTAC 2379
Db 2281 CTAGCCTTCTATCAATCAACATCAAGTGAATATTTTACTAAATAATGATATCTCATATTAC 2340
Qy 2380 ACGAGTAAATAGACATAAATAAAGTGGAGCACATTTAAGTAAATATTAATAAATTAAGTCAG 2439
Db 2341 ACGAGTAAATAGATTAATAAATAAAGTGGAGCACATTTAAGTAAATATTAATAAATTAAGTCAG 2400
Qy 2440 TTAATCTAAATTTGATACATCATTTTTCATATCATGACAGCCATTTGTTAAATAAT 2499
Db 2401 TTAATCTAAATTTGATACATCATTTTTCATATCATGACAGCCATTTGTTAAATAAT 2460
Qy 2500 GACAGCTATGCTTATATGAAAAAATATGATCGGCATGAATTTCTCAGCATTAACACAT 2559
Db 2461 GACAGCTATGCTTATATGAAAAAATATGATCGGCATGAATTTCTCAGCATTAACACAT 2520
Qy 2560 GATTGGATCGAGAAATCAATGCGCATCCACCATTTAAAAAGCTGATTTAAAAACCTATTTT 2619
Db 2521 GATTGGATCGAGAAATCAATGCGCATCCACCATTTAAAAAGCTGATTTAAAAACCTATTTT 2580
Qy 2620 AATGACATGACTTAAAGAAATGATGAATGTAAGGGGCATCACAAAGTATGTTTATGAAG 2679
Db 2581 AATGACATGACTTAAAGAAATGATGAATGTAAGGGGCATCACAAAGTATGTTTATGAAG 2640
Qy 2680 TATGCGTACCGCATGAGCTTCTGACGATTTATTAAGAAAGTCAATCAGCATCTGCCAATCA 2739
Db 2641 TATGCGTACCGCATGAGCTTCTGACGATTTATTAAGAAAGTCAATCAGCATCTGCCAATCA 2700
Qy 2740 ATTGATAGTGGCCAGAAATATAACATGAGGATATTTGGTTCCCAATTTGCATTTTAAATC 2799
Db 2701 ATTGATAGTGGCCAGAAATATAACATGAGGATATTTGGTTCCCAATTTGCATTTTAAATC 2760
Qy 2800 TTAGAAAGAAACCGCCATGATTTTAAATAAATCAATCGACCTGATATGCTTGG 2859
Db 2761 TTAGAAAGAAACCGCCATGATTTTAAATAAATCAATCGACCTGATATGCTTGG 2820
Qy 2860 GAACGAAATTTACAAATGACAAATGAACAAATTCAAAGTGCAGAAAGGAGGAAATATC 2919
Db 2821 GAACGAAATTTACAAATGACAAATGAACAAATTTGAAAGTGCAGAAAGGAGGAAATATC 2880
Qy 2920 CCCGTTAAAGTTCATTTTAAATAGTAAACGGTATATAA 2959
Db 2881 CCTGTTAAAGTTCATTTTAAATAGTAAACCTCTATAA 2920

RESULT 6

US-10-216-289-3
; Sequence 3, Application US/10216289
; GENERAL INFORMATION:
; APPLICANT: NINOMIYA, TOSHIO
; APPLICANT: SUGIURA, NOBUO
; APPLICANT: KIMATA, KOJI
; TITLE OF INVENTION: CHONDROITIN POLYMERASE AND DNA ENCODING THE SAME
; FILE REFERENCE: 226882USO
; CURRENT APPLICATION NUMBER: US/10/216.289
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: JP 2001-244685
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: JP 2001-324127
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: JP 2002-103136
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 14483
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3787)..(5847)

; OTHER INFORMATION:
US-10-216-289-3
Query Match 27.3%; Score 814.2; DB 6; Length 14483;
Best Local Similarity 63.0%; Pred. No. 1.6e-146;
Matches 1312; Conservative 0; Mismatches 758; Indels 13; Gaps 3;
Qy 53 GAAAGAAATGAATACATTTATCACAAGCAATAAAAAGCATATAACAGCAATGACTATGAAT 112
Db 3779 GAAATAAATGAGTATTTCTTAATCAAGCAATAAATTTATATAAACAATAAATATCGCC 3838
Qy 113 TAGCACTCAAAATATTTGAGAAAGTCTGCGAAACCTACGGCGAAAAATTCGTTGAATTC 172
Db 3839 AAGCTTTATCTCTTTTGAAGAGGTTGCGAAATTTATGATGTTAGTGGTCGGAAGCAA 3898
Qy 173 AAATTTATCAAAATGAAGAAACCTCTGACCAATTTCTTATGTAAGTGAAGATAAAAAA 232
Db 3899 ATATAAAATTTATGCCAACCCGCACT-----CAATCTTTCTGAAGAAGTTGATAAGTTAA 3952
Qy 233 ACAGTGTGTTGCGATAGCTCATTTAGATATGCGAACACAGCTCTTACTTTTCCACGTAATAA 292
Db 3953 ATCGTAAAGCTGTTATTTGATATTTGATGCGAGCAACAAAAATAATGTTCTTAAGCCCAAG 4012
Qy 293 AATTAACTCTATCCGAATCAGAAAAACAGTTTAAAAAATTAATGGAATATCTATCACTG 352
Db 4013 CAATTAGTCTGAACGAGTTGAAAAAATAAATAAGCAAAATACCGAGAAATAACCG 4072
Qy 353 GGAATAAATCGAGAACGCGAGAAATCAGAAAGTGGAACTAGTACCCAAAGATTTTCCCTA 412
Db 4073 CAAGAATATCAGAACGGCGGAGTTAAAGAAAGTGAACCCATTCCTTTAGATTGGCCTA 4132
Qy 413 AAGATCTTTGTTGCTCCATTTGCCAGATCATGTTAATGATTTTACATGGTACAAAAATC 472
Db 4133 GTGATTTAACTTTACCGCGTTTACCTGAGAGCAACAAAGATTATGTTTGGCGGGGAAAA 4192
Qy 473 GAAAAA-----AAGCTTAGGTATAAAGCTGTAAATAAGAAATATCGGTCCTTTCTATTAT 527
Db 4193 GAAAAAGAGCTTGATGATTTATCCAAGAAAAACAGTTAAT--CATTTGACGGCTTAGTATTGT 4250
Qy 528 TATTCCTACATTTAATCGTAGCGGTATTTTATAGATAAAGCTTAGCTGTTTGGTCAATCA 587
Db 4251 AATTCCTACATATAATCGAGCAAAATACTTCCAATTACACTGCTGTTTGTAAACCA 4310
Qy 588 GAAAAAACTACCCATTTGAAAGTCTGTTGTCAGATGATGTTAGTAAAGAAACTTACT 647
Db 4311 AAAGACCATATACGACTATGAAGTTATCTTCCGATGATGGAAGTAAAGAAATATGA 4370
Qy 648 TACCATTGTGCAAAATATAGCAAAACCTTGACATAAAAGTATGTAAGACAAAAAGATT 707
Db 4371 AGAAATAGTAAGAGAAATTTGAAAGTTTATTAATAATAAATAATATGACGTCAGAGGATTA 4430
Qy 708 TGGATATCAATTTGTCGAGTCAGAACTTAGCTTTACGTACAGCAAGTATGATTTTGT 767
Db 4431 TGGATATCACTGTCGTCGTTAGAAATCTTGGCTTAGGCTGCAAGATTAATATTATG 4490
Qy 768 CTCGATTTCTAGACTGCGATATGCGACCAACAATAATATGGTTTCAATCTTATCTTACAGA 827
Db 4491 TCGAATCTCGGATTTGATATGCTCCGAAACCCACTATGGTTTCACTATATATGAAC 4550
Qy 828 ACTATTAGAACAAATGATATTGTTTAAATTTGACCTAGAAAAATATGTTGATCTATCAA 887
Db 4551 ATTAGCGTGGACGATTAATGTTCTCTAATTTGGCCCTAGAAAAATATATAGATACAAAGCAA 4610
Qy 888 TATTACCGCAGAACCAATTCCTTAACGATCCATTTAATTAAGATAATCACTACCTCAAAACCGC 947
Db 4611 GCATACATATTAGATTTCCTTTCCCAAAATCACTAATAAATGAATTTCTCGAAATCAT 4670
Qy 948 TACAATAAATCTCTTCGATTACATCAAAAGGAAATATATCGTTGGATTGGAGATTAGA 1007
Db 4671 TACTAATAATCAGTTTGCAGCAAGGTTGAGCAAAACAAATCAGTTTGACTGGCGAATAGA 4730
Qy 1008 ACATTTCAAAAAAACCGATAATCTACGCTATGTTGATTTCTCCGTTTCTGTTTCTTTCGTC 1067

Db 4731 ACATTTCAAAAATACCGATAATCTAAGATTATGCAACACACCATTCGATTTTTCAGCGG 4790

Qy 1068 GGGTAATGTTGCATTTCTPAAAGAATGGCTAAATAAAGTAGGTGGTTCGATGAAGAATT 1127

Db 4791 AGGTAATGTCGTTTTCGCAAAAATGGCTTTCCGTGCGAGGATGGTTTGTGATGAAGAGTT 4850

Qy 1128 TAATCAATGGGGGGGGAAGATGTAGAATTTGGTTACAGATTAATTTGCCAAAGGCTGTTT 1187

Db 4851 TAGCCATTTGGGGGGGAGGATAATGAGTTGGATATCTCTCTACAGAGAGGATGTTA 4910

Qy 1188 TTTCAGAGTAATTTGACGGCGGAATGCCATCCATCAAGAACACACCTGGTAAAGAAAATGA 1247

Db 4911 CTTTCGGTCTGTGAAGGAGCATGGCATATCATCAAGAACACACCGCGGGAAGAAAACGA 4970

Qy 1248 AACAGAACCGAAGCTGGTAAAGATTAATCGCTPAAATTTGTGAAAGAAAGTACCTTA 1307

Db 4971 GACGGATCGTGGCGGAGGAAAATATTAAGTTCTCAATTTGTTACAGCAAAAAGTTCCTTA 5030

Qy 1308 CATCTATAGAAAGCTTTTACCATAATGAAGATTCCACATATTCATAGAAATACCTTTAGTTTC 1367

Db 5031 TTTCATAGAAAAGAAAAGAAAATAGAACCGGACATTAAGAAAGAGTACCACCTAGTATC 5090

Qy 1368 TATTTATATCCCGCTTATAACTGTGCAATTAATTAATCAAGATGTGTAGATAGTGTCT 1427

Db 5091 TATATATATCCCGCTTATAACTGTCTAAATATATTTGTTGTTGTTGAAAAGCGCCCT 5150

Qy 1428 TAATCAAACTGTTGTCGATCTCGAGTTTGTATTTGTAACGATGGTTCACAGATATATAC 1487

Db 5151 TAATCAGACAAATAACTGACTTAGAAGTATGCATATGGCATGTGGTTTCCACAGATGATAC 5210

Qy 1488 CTTAGAAAGTGATCAATAAGCTTTTGTGTAATAATCTTAGGTAGCGCATCATGTAAACC 1547

Db 5211 ATTCCGAGTCTTCAGGAGCATATGCAACACCATCTCGAGTTGTTTATTTTACACAAA 5270

Qy 1548 AAATGGCGGAATAGCCTCAGATCAAAATGACGCGGTTTCTTTTCTTAAAGGTTATTACAT 1607

Db 5271 AAACAAAGGAATGGTTGACGATCTAATACAGCAGTATAGATTGTTGCGGGATTCATAT 5330

Qy 1608 TGGCAGCTAGATTCAGATGATTTCTGAGCCGTGATGCACTGTAACCTGTTTAAAAGA 1667

Db 5331 AGGTCAGTTAGACTCTGATGACTTCTTGAACAGATGCTGTTGCACTATGTCTAGATGA 5390

Qy 1668 ATTTTAAAAGATAAACCGCTAGCTTGTGTTTATACCACTAATAGAACGTCATTCGGA 1727

Db 5391 ATTTAGAAAAGATCTATCATTTGGCATGTGTTTATACAACTAACCGTATATAGATCGTGA 5450

Qy 1728 TGGTAGCTTAATCGCTAATGTTTCAATTTGCCAGATTTTTCACGAGAAAACCTCACAC 1787

Db 5451 AGGTAATTTGATATCAAAATGGCTATAATGGCCCATTTTATTCGGGAGAAAACCTACTAG 5510

Qy 1788 GGTATGATTCCTCACCACTTTTAGAATGTTTACGATTTAGAGCTTTGGCATTTAACGGATGG 1847

Db 5511 TGCAATGATATGTCATCTATTCAGGATGTTTCAACAGAGAGCATGGAACCTTAACCTGAAG 5570

Qy 1848 ATTTAACGAAATATTTGAACCGCGTGGATATGACATGTTTCTTAACTCAGTGAAGT 1907

Db 5571 TTTCACGAATCGATFACGACGAGTGGATTACGATATGATTAAACCTTAGTGAAGT 5630

Qy 1908 TGGAAAATTTAAACATCTTAATAAATCTGCTATAACCGCGTATTACATGTTGATGAACAC 1967

Db 5631 TGGACCGTTCAAGCATATAACAAATTTGTTATTAATCCGGTATTGCATGGTGAATAATAC 5690

Qy 1968 ATCCATTAAAGAACTCGGCATTTCAAAAGAAAACCATTTTGTGTAGTCAATCAGTCAAT 2027

Db 5691 GTCTATAAAAAGTTGGATATTTCAAAAGGAAAATCATTTTAAAGTTGTTAAAGAAATCAT 5750

Qy 2028 AAATAGACAGGCATCAATTAATAATATGACAAATTTGATGATTTAGATGAAGATAG 2087

Db 5751 AAGTAGCTAGCATATAAAAATATAAATATTTCACCATTAACATTAATTTGAATGAATGATG 5810

Qy 2088 AAGTATATCTTCAATAAACCGCTGAATATATCAAGAAAGAAATG 2130

Db 5811 AAAATATCTCGGAAAATAATAGAGAAATGATTTATAATATTG 5853

RESULT 7

US-10-216-289-1

; Sequence 1, Application US/10216289

; GENERAL INFORMATION:

; APPLICANT: NINOMIYA, TOSHIO

; APPLICANT: SUGIURA, NOBUO

; APPLICANT: KIMATA, KOJI

; TITLE OF INVENTION: CHONDROITIN POLYMERASE AND DNA ENCODING THE SAME

; FILE REFERENCE: 226882050

; CURRENT APPLICATION NUMBER: US/10/216,289

; CURRENT FILING DATE: 2002-11-04

; PRIOR APPLICATION NUMBER: JP 2001-244685

; PRIOR FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: JP 2001-324127

; PRIOR FILING DATE: 2001-10-22

; PRIOR APPLICATION NUMBER: JP 2002-103136

; PRIOR FILING DATE: 2002-04-05

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 2058

; TYPE: DNA

; ORGANISM: Escherichia coli

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(2058)

; OTHER INFORMATION:

US-10-216-289-1

Query Match 27.1%; Score 807.2; DB 6; Length 2058;

Best Local Similarity 63.2%; Pred. No. 3.le-145;

Matches 1296; Conservative 0; Mismatches 743; Indels 13; Gaps 3;

Qy 61 ATGATACATTTATCACAGCAATAAAGCATATATACAGCAATGACTATGAATTAGCACTC 120

Db 1 ATGAGTATCTTATCAAGCAATAAATTTATATAAAACAAAATTTATGCGCAAGCTTTA 60

Qy 121 AAATTTATTTGAGAAGTCGTGAAACCTTACGGCGGAAAATCGTTGAATTCCAAATATC 180

Db 61 TCTCTTTTGGAAAGGTTGCTGAAATTTATGATGTTAGTTGGGTGCGAAGCAAAATATAA 120

Qy 181 AAATGTAAGAAAGAACTCTCGACCAATTCCTTATGTAAGTGAAGATAAAAAACAGTGT 240

Db 121 TTATGCCAACCGCACT-----CAATCTTCTGAAGAGTTGATGAATTAATCGTAA 174

Qy 241 TCGATAGCTCATTTAGATATCGCAACACAGCTCTTTACTTTTCCACGTAACCAATTAAC 300

Db 175 GCTGTTATTGATATTTGATGTCAGCAACAAAATAATGTTCTTACGCCAAAGCAATTAG 234

Qy 301 CTATCCGAATCAGAAAAAAACAGTTTAAAAATAAATGAAATCTATCACTGGGAAAAAA 360

Db 235 CTGAACGAGGTTGAAAAAATAAATAAAGCAATACCGAATAACCGAATAACCGAAGAA 294

Qy 361 TCGGAGAACGCAAGAAATCAGAAAGGTGGAACTAGTACCCAAAGATTTTCTTAAAGATCT 420

Db 295 TCAGAACGGCGGAGTTAAGGAGTCGAAACCCATTCCTTTAGATTGGCTAGTATTGA 354

Qy 421 GTTCTGTCTCAATGCGCAGATCATGTTAATGATTTTACATGTTGACAAAATCGAAAAA- 479

Db 355 ACTTTACCGCCGCTTACCTGAGAGCAACAAACGATTAATTTGTTGGCGGGGAAAGAG 414

Qy 480 ----AAGCTTAGGTATAAGCCTGTAATAAAGAAATCGTCTTTCTATTATTATTCCTA 535

Db 415 CTTGATGATTTATCCAGAAAAACAGTTAAT--CAATCAGCGGCTTAGTATTGTAATTCCTA 472

Qy 536 CATTTAATCGTAGCCGCTATTTTAGATATATAACGTTAGCCTGTTTGGTCAATCAGAAAA 595

Db 473 CATATAATCGCAAAAATACCTTGCAATTCACCTTCTGCTTCTTCTGTAACCAAGACCA 532

Qy 596 ACTACCACTTTGAAGTCGTTGTTGAGATGATGTTAGTAAAGAAACTTACTTACCATTG 655

Db	533	TATACGACTATGAAGTTATTGTTGGCGATGATGGGAAGTAAAGAAAAATATTGAAGAAATAG	532
Qy	656	TGCAAAAATACGCAACAAAAACCTTGACATAAAGATATGTAAAGACAAAAAGATTATTCGATATC	715
Db	593	TAAGAGAATTGAAAGCTTTTATTAAATATATAAATATGATACGTCAGAGAGGATTATGATATC	652
Qy	716	AATTGTGTGCGAGTCAGAAACCTTAGGTTTACGTCACGACAAAGATATGATTTGTCTCTCGATTC	775
Db	653	AACGTGTGCTGTGTAGAAATCTTTGGGCTTTAGGGCTGCAAAAGTATAAATTATGTTGCGAAATC	712
Qy	776	TAGACTGCGATATGGCACCACACAAATATATGGTTTCATCTTATCTTACAGAACATTTAG	835
Db	713	TGGATTGTGATATGGCTCCGACCCCACTATATGGTTTCAGTCATATATGGAACATTTAGCGG	772
Qy	836	AAGACAATGATATTTGTTTTAAATTGGACCTTAGAAAAATATGTGGATACTACTAATAATTACCG	895
Db	773	TGGACGATATGTTGCTCTAATTGGCCCTAGAAAATATATAGATACAGCAAGCATACAT	832
Qy	896	CAGACAATTCCTTTACGATCCATATTTAATAGAAATCACTACCTGGAACCGCTACAATA	955
Db	833	ATTTAGATTTCCTTTCCCAAAAATCACTAATAAATGAAATTCCTGAAATCATTTACTAATA	892
Qy	956	ACAATCCTTCGATTACATCAAAAGGAATATATCGTTGGATTGGAGATTAGAACATTTTCA	1015
Db	893	ATCAGGTTGCAGCGAAGTTGAGCAAAAACAACTAGTTGACTGGCGAATAGAACATTTTCA	952
Qy	1016	AAAAACCGATAATCTACGCTATGTGATCTCCGCTTTCCGTTATTTTGTCTCGGGTAATG	1075
Db	953	AAAATACCGATATCTAAGATTTATGCACACACCAATTTTCGATTTTGTAGCGGAGGTAA	1012
Qy	1076	TTGCATTTTCTAAAGATGGCTAAATAAAGTAGTGGTTTCGATCAAGAAATTTTAATCATTT	1135
Db	1013	TCGCTTTTGGCAAAAAATGGCTTTTCCGTGCGAGGATGGTTTGATGAAGAGTTTACGCATT	1072
Qy	1136	GGGGGGCGAGATGTAGAAATTTGGTTACAGATTTATTTGCCAAAGCGTGTTTTTTCAGAG	1195
Db	1073	GGGGGGGGAGGATAATGAGTTTGGATATCGCTCTACAGAGAAGGATGTACTTTTCGGT	1132
Qy	1196	TAAATTGACGGCGGAATGGCCATCCATCAAGAACCACTGGTTAAAGAAAAATGAACACAAC	1255
Db	1133	CTGTTGAGGAGCAATGGCAATATCATCAGAACCAACCCGGGAAAGAAACGACAGCGATC	1192
Qy	1256	CGGAAGCTGTAAAGATTTACGCTTTAAATTTGTGAAAGAAAAGGTACCTTTACATCTATA	1315
Db	1193	GTGCGCGAGGAAAAATATTACTGTTCATTTGTTACAAATGTTTACAGCAAAAAGTTCCTTATTCTATA	1252
Qy	1316	GAAAGCTTTTACCAATAGAGATTACATATTCATAGAAATACCTTTAGTTTCTATTTTATA	1375
Db	1253	GAIAAAAAAGAAAAATAGAAATCCGGCGACATTTAAAAGAGTACCACTAGTATCTATATATA	1312
Qy	1376	TCGCCGCTTATAACTCTGCAAAATATATTCAAGATGTGTAGATAGTCTCTTAATCAAA	1435
Db	1313	TTCCCGCTATACTGCTCTAATATATTGTTTCGTTGTGTTGAAAGCCCTTAAATCAGA	1372
Qy	1436	CTGTGTTCGATCTCGAGGTTGTATTGTTGAACGATGGTTCAACAGATAAATACCTTTAGAAG	1495
Db	1373	CAATAACTGACTTTAGAAGTATGCATATGCGATGATGGTTCCACAGATGATACATTTGCGGA	1432
Qy	1496	TGATCAATAGCTTTTATGTAAATATCTCTAGGGTAGGCATCATGTCTTAAACCAATGGCG	1555
Db	1433	TTCTTTCAGGAGCATATGCAAAACCACTTCCTCGAGTTTCGTTTTTATTTCACAAAAAACAAG	1492
Qy	1556	GAATACCTTCAGCATCAAAATGCAGCGGTTCTTTTGGCTTAAAGGTTATTACATTCGGGCAGT	1615
Db	1493	GAATTGTTCCAGCATCTAATACAGCAATTTAGATTGTGTGCGGGGATCTTATATAGGTCAGT	1552
Qy	1616	TAGATTTCAGATGATTATCTTGAGCCTGATGCAAGTTTGAACCTGTGTTTAAAAAGAAATTTTAA	1675
Db	1553	TAGACTCTGATGACTTCTTTGAACCAAGATGCTGTTTGAACATGTCTAGATGAATTTAGAA	1612
Qy	1676	AGATAAAGCGTAGCTTGTTTATACCACTAATATAGAAACGTCOAATCCCGATCGGTAGCT	1735
Db	1613	AGATCTCATTTGGCATGTGTTTATACAACTTAAACCGCTTAATATAGATCGTGAAGGTAATT	1672

QY 1736 TAATCGCTAANTGGTTACAAATTGGCCAGCAAGAATTTTTCAGGAGAAAACACTCACAAACGGCTATGA 1795
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1673 TGATATCAAATGGCTATATAATTGGCCCCATTTATTTCGCCGAGAAAAACCTTACTAGTGCAATGA 1732

QY 1796 TTGCTCACCATTTTAGAATGTTTTCAGATTAGAGCTTTGGGCATTTTAACGGATGGATTTAACG 1855
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1733 TATGTCATCAITTTTCAGGATGTTTCACAGAAAGAGCATGGAAACCTNACTGAAGGTTTCCAACG 1792

QY 1856 AAAATATTGAAACCCGCTGGATTATGACATGTTCCCTTAAACCTCAGTGAAGTTGGAAAAAT 1915
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1793 AATCGATCAGCAACGACGTTGATTACGATATGATTTTAAAACTTAGTGAAGTTGGACCGT 1852

QY 1916 TTAACATCTTAATAAATCTCCATACCCGGGTATTACATGGTGATACACATCCATT 1975
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1853 TCAAGCATATAACAATAATTTGTTTAAATCCGGTATTGCATGGTGAATAACGTCTATAA 1912

QY 1976 AGAAACTCGGCATTCAAAGAAAAACCAATTTTGTGTACTCAATCAGTCATTAATAATAGAC 2035
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1913 AAAGTTGGATATTCAAAAGGAAATCATTTTTAAAGTTGTTAACGNATCATTAAGTAGGC 1972

QY 2036 AAGGCATCAATPATTATTAATTACAAATTTTGATGATTTAGATGAAAGTAGAAAGTATA 2095
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1973 TAGGCATAAAAAATATAAATATTCACCATTAACTTAATTTGAATGAATGTGAAAAATATA 2032

QY 2096 TCCTCAATAAAA 2107
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2033 COTGGGAAAAA 2044

RESULT 8
US-09-134-000C-2987
; Sequence 2987, Application US/09134000C
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2987
; LENGTH: 993
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-2987

	Query Match	2.9%;	Score 86;	DB 5;	Length 993;	
	Best Local Similarity	54.0%;	Prid. No.	8e-08;		
	Matches 176;	Conservative	0;	Mismatches 150;	Indels	Gaps
QY	1354	ATACCTTTAGTTCCTATTATATCCCCGCTTATAACTGTGCRAATTATATTTCAAAAGATGT	1413			
Db	19	ATGCCCAAATTAGTATTATTCTCCTGTATACAATGTAGAAAAAATATTAGAAAATGT	78			
QY	1414	FPAGATAGTGCTCTTAATCCAAACTGTGTGCAGTCTCGAGGTTTGTTGTTTGTACACGATGGT	1473			
Db	79	GPACGCTCTATTTTTAGCTCAAAAGCTTTACTGACTTTTGAATTAATCTCGTGGACGATGGC	138			
QY	1474	TCAACAGATAATACCTTTAGAAGTGATCAATAAGCTTTATGGTAAATAATCCTAGGGTACGC	1533			
Db	139	TCTCAGACAGTCTGGAGCATGTGTATCAGTTTGTCTGAACAAGATCAACGGGTAAAA	198			
QY	1534	ATCATGTCTTAATCCAAATGGCGGAATAGCCTCAGCATCAAAATCAGCCGCTTTCTTTTGGT	1593			
Db	199	GTTATCCATAAGAAAAATSGTGGCTAACCGATGCTCGTAAATGCTGGAATTGAATAGCA	258			
QY	1594	AAAGGTTATTACATTTGGGCAGTTAGATTCAGATGATATTCTTTGAGCCTGATCAGTTGAA	1653			
Db	259	ACAGGTGAGTATTTAGTTTTCTAGATAGTATGATTACATTTGCAGATGATATGTATGAA	318			

QY 1654 CTGTGTTTAAAGAAATTTTAAAGCA 1679
|| | | | | | | | | | | | | | | | | | | | |
Db 319 CTATTATACAAATATCGTAAAGCA 344

RESULT 9

US-10-303-161-1/c
; Sequence 1, Application US/10303161
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,161
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus
US-10-303-161-1

Query Match 2.7%; Score 79.2; DB 6; Length 11474;
Best Local Similarity 51.2%; Pred. No. 1.8e-06;
Matches 211; Conservative 0; Mismatches 198; Indels 3; Gaps 1;

QY 1296 AAAGGTACCTTACATCTATAGAAAGCTTTTACCAATAGAGATTCACATATTCATAGAAT 1355
|| | | | | | | | | | | | | | | | | | | | |
Db 11424 AAATTCCTAAACTTTGATAGAATGTTTATATTTTAAACAAATTTAGGAAATAT 11365
|| | | | | | | | | | | | | | | | | | | | |
QY 1356 ACCTTTAGTTCTTATATATCCCGCTTATACTGTGCAAAATTTATTTCAAGATGTGT 1415
|| | | | | | | | | | | | | | | | | | | | |
Db 11364 GCCACAACTTCTATCATATATCCCGCTTTTAAATCTTGGCATTTATCTCAAGAGCTTT 11305
|| | | | | | | | | | | | | | | | | | | | |
QY 1416 AGATAGTCTCTTAATCAAACTGTTCGATCTCGAGTTTGTATTTGTACGATGGTTC 1475
|| | | | | | | | | | | | | | | | | | | | |
Db 11304 GCAAAAGTTGTATAATCAGACTTTTAAAGATATTGAAATTTGATATTCGATGATAAAG 11245
|| | | | | | | | | | | | | | | | | | | | |
QY 1476 AACAGATAATACCTTAGAAGTGATCAATAAGCTTTATGTTATATCTAGGGTACGCAT 1535
|| | | | | | | | | | | | | | | | | | | | |
Db 11244 TAAAGATAATAGTTTAAATAGTTTGTAGATTTTGCACCAATTAAGATTAAGAT 11185
|| | | | | | | | | | | | | | | | | | | | |
QY 1536 CAT---GTCTAAACCAATGCGGAATAGCCTCAGCATCAATCGACCGTTCTTTTTCG 1592
|| | | | | | | | | | | | | | | | | | | | |
Db 11184 CTTTCAAAATGAAGAAATTTAGGCACCTTTTGCAGATGAGAAATTTGGGGGTTTTCGATTC 11125
|| | | | | | | | | | | | | | | | | | | | |
QY 1593 TAAAGTTATTACATTTGGGCAGTTAGATTCAGATGATTATCTTGAGCCTGATGAGTTGA 1652
|| | | | | | | | | | | | | | | | | | | | |
Db 11124 TAGTCTGATTTTAAATGTTTGTAGATGATGATTTTAAACACCTGATGCTTGGA 11065
|| | | | | | | | | | | | | | | | | | | | |
QY 1653 ACTGTGTTTAAAGAAATTTTAAAGATAAAACCGCTAGCTTGTGTTTATACC 1704
|| | | | | | | | | | | | | | | | | | | | |
Db 11064 AATAGCGTTTAAAGAGATGAAAAAGGTTTGAATTTGCTTTGTTTGATGCC 11013
|| | | | | | | | | | | | | | | | | | | | |

RESULT 10

US-10-303-162-1/c
; Sequence 1, Application US/10303162
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel

; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus
US-10-303-162-1

Query Match 2.7%; Score 79.2; DB 6; Length 11474;
Best Local Similarity 51.2%; Pred. No. 1.8e-06;
Matches 211; Conservative 0; Mismatches 198; Indels 3; Gaps 1;

QY 1296 AAAGGTACCTTACATCTATAGAAAGCTTTTACCAATAGAGATTCACATATTCATAGAAT 1355
|| | | | | | | | | | | | | | | | | | | | |
Db 11424 AAATTCCTAAACTTTGATAGAATGTTTATATTTTAAACAAATTTAGGAAATAT 11365
|| | | | | | | | | | | | | | | | | | | | |
QY 1356 ACCTTTAGTTCTTATATATCCCGCTTATACTGTGCAAAATTTATTTCAAGATGTGT 1415
|| | | | | | | | | | | | | | | | | | | | |
Db 11364 GCCACAACTTCTATCATATATCCCGCTTTTAAATCTTGGCATTTATCTCAAGAGCTTT 11305
|| | | | | | | | | | | | | | | | | | | | |
QY 1416 AGATAGTCTCTTAATCAAACTGTTCGATCTCGAGTTTGTATTTGTACGATGGTTC 1475
|| | | | | | | | | | | | | | | | | | | | |
Db 11304 GCAAAAGTTGTATAATCAGACTTTTAAAGATATTGAAATTTGATATTCGATGATAAAG 11245
|| | | | | | | | | | | | | | | | | | | | |
QY 1476 AACAGATAATACCTTAGAAGTGATCAATAAGCTTTATGTTATATCTAGGGTACGCAT 1535
|| | | | | | | | | | | | | | | | | | | | |
Db 11244 TAAAGATAATAGTTTAAATAGTTTGTAGATTTTGCACCAATTAAGATTAAGAT 11185
|| | | | | | | | | | | | | | | | | | | | |
QY 1536 CAT---GTCTAAACCAATGCGGAATAGCCTCAGCATCAATCGACCGTTCTTTTTCG 1592
|| | | | | | | | | | | | | | | | | | | | |
Db 11184 CTTTCAAAATGAAGAAATTTAGGCACCTTTTGCAGATGAGAAATTTGGGGGTTTTCGATTC 11125
|| | | | | | | | | | | | | | | | | | | | |
QY 1593 TAAAGTTATTACATTTGGGCAGTTAGATTCAGATGATTATCTTGAGCCTGATGAGTTGA 1652
|| | | | | | | | | | | | | | | | | | | | |
Db 11124 TAGTCTGATTTTAAATGTTTGTAGATGATGATTTTAAACACCTGATGCTTGGA 11065
|| | | | | | | | | | | | | | | | | | | | |
QY 1653 ACTGTGTTTAAAGAAATTTTAAAGATAAAACCGCTAGCTTGTGTTTATACC 1704
|| | | | | | | | | | | | | | | | | | | | |
Db 11064 AATAGCGTTTAAAGAGATGAAAAAGGTTTGAATTTGCTTTGTTTGATGCC 11013
|| | | | | | | | | | | | | | | | | | | | |

RESULT 11

US-10-303-118-1/c
; Sequence 1, Application US/10303118
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,118
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01

Db	11364	GCACAACCTTCTATCATATAATCCCGCTTTTAAATCTTGGATTTATCTCAACAGCTTT	11305
Qy	1416	AGATAGTGCCTTAATCAAACTGTTGTGCAGTCTCGAGGTTTGTAATTTGTAACGATGTTTC	1475
Db	11304	CCAAGTGTATAAATCAGACTTTAAAAGATATTGAAATTTTGATTTATCGATGATAAAAG	11245
Qy	1476	AACAGATAAATACCTTAGAAGTGATCAATAAGCTTTATGGTAAATAATCCTAGGTACGCAT	1535
Db	11244	TAAAGATAAATAGTTTAAATATGGTTTTAGAAATTTGCAAAAAGATCCAGAATAAAAT	11185
Qy	1536	CAT---GTCTAAACCAATGCGGNAATAGCCTCAGCATCAATCAGCGTTTCTTTTGC	1592
Db	11184	CTTTCAAAATGAAGAAAAATTTAGGCACCTTTGCAAGTAGAAATTTGGGGGTTTTGCAATC	11125
Qy	1593	TAAAGGTTATTACATTTGGCGCAGTTAGATTACAGATGATTCTTGAGCCTGATGCAGTTGA	1652
Db	11124	TAGTTCTGATTTTATAATGTTTTTAGATAGTGATGATTTTTTAACACCTGATGCTTGGA	11065
Qy	1653	ACTGTGTTTAAAGAAATTTTTAAAAGATAAACCGCTAGCTTGTTGTTATACC	1704
Db	11064	ATATGCGTTTTAAAGAGATGAAAAAGGTTTTGATTTGCTTTGTTTTGATGCC	11013

RESULT 14

```

RES001 14
US-10-240-454-27/c
; Sequence 27, Application US/10240454
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Metabolism
; FILE REFERENCE: 5013.1010
; CURRENT APPLICATION NUMBER: US/10/240,454
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/04016
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 68
; SEQ ID NO 27
; LENGTH: 17527
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-454-27

```

	Query Match	2.6%;	Score 77.2;	DB 6;	Length 17527;
	Best Local Similarity	43.9%;	Prod. No. 4.3e-06;		
	Matches 434;	Conservative	0;	Mismatches 543;	Indels 11; Gaps 2;
QY	1992	AAAGAAAAACCACTTTTGTGTAGTCAATCAGTCATTAAATAGACAAGGCATCAATATTATTA	2051		
Db	16996	ATAAAAAANACGTATCTTAATACTTTAAAAACATTTAAAAAATAACTAAACCACACGA	16937		
QY	2052	TAAATATGACAAATTTGATGATTTAGATGAAGTAGAAGTATATCTTCAATTAANAACGC	2111		
Db	16936	TAAATCAACCTAAAAATCCCAACTACTCAAAAAACAAAAATAAATATTATTACTTTAAACCC	16877		
QY	2112	TGAATATCAAGAGAAGTAATGGATATGTTAAAGATCTTTAACTCATTTCAAAATTAAGATGC	2171		
Db	16876	AAATTTAAATTAACAAATAAACCAAAATCACACCCTACTCTCAACCTTAAACAACAAA	16817		
QY	2172	CAAAATCGGAGTCAGTATTTTCTATCCCAATATCAATTAACGGCTTAGTCAAAAACATAA	2231		
Db	16816	CAAAACCCCACTTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAATCA	16757		

Qy	2232	CAATATTATTGAATATAATAAAAAATATATTTCGTATTATTCTCATAGTTGATPAAGAATCA	2291
Dd	16756	AAAAAAAAACAAACAAAACAAAAAATCTTCCCAATAAT---CAGTATTAAACTFAAA	16701
Qy	2292	TCTTACACCAGACATCAAAAAGAAATATTGGCTTCTATCATTAAGCACCAAGTGAAAT	2351
Dd	16700	AAAAAAAAATAAAAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC	16641
Qy	2352	TTTACTAATAATGACATCTCATATTACACAGTAGTAATAGACTAATAAAAACTGAGGCACA	2411
Dd	16640	TTTTCAAATAATAACGACTCAATATAAAAAACAATAATTTCTTTATACAAC-----AT	16588
Qy	2412	TTTTAAGTAATATTATAAATTTAAGTCAGTTAAATCTAAATTTGTAATACATATTTTGA	2471
Dd	16587	TTTTTAATTAACCTAAAAATTTATACAACATTTTTTAAAAAATCGAACCACTCAATCA	16528
Qy	2472	TAATCATGACAGCCATTTCGTTTAAAAATGACAGCTATGCTTATATGAAAAAATATGATG	2531
Dd	16527	TACCTATAATCCTAACACTTTTAAAAAATCAAAACAAATAAATCACCTFAAAATCAAAAT	16468
Qy	2532	CGCGATGAATTCCTCAGCATTAACACACATGATTGGATCGAGAAATCAATGGCGATCCACC	2591
Dd	16467	TACGACCAACAACTAACCAAAATTAACGAAATCCGGTCTCTACTAAAAACACACAAAAAA	16408
Qy	2592	AATTTAAAAAGCTGATTAACACCTATTTTAAATGACAATGACTTTAAGAGTATGAATGTGAA	2651
Dd	16407	TTAACGAACATATAATACACACACACTATAATCCCACACTACTCAAAAAACTAAAAAAT	16348
Qy	2652	AGGGGATCACAAGTGATTGTTGAAGTATGCCGTACCGCATGAGCTTCTGACGATAT	2711
Dd	16347	TACTTAAAGCCAAAAAATAAAAAATTTACTATAAACCTTAATCCTACCATTAACCTCCAACC	16288
Qy	2712	TAAAGAGTCATCACATCCTCGCAATCAATTGATAGTGTGCAGCAATATAACACTGAGGA	2771
Dd	16287	TAAACAAAAAACGAAACTCCTCAAAAAAATTAATTAATTTAATTAATTTAAAAA	16228
Qy	2772	TATTTGGTTCCAAATTTGCACCTTTTAATCTTAGAAAAGAAACCGGCCATGTATTAATAA	2831
Dd	16227	ATATAATCTCTATATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAA	16168
Qy	2832	AACATCGACCTGACTTATATGCCITGGGAACGAAATTTACAAATCGACAAATGAACAAAT	2891
Dd	16167	AACCTTTACCAATAATAAAACCAAAAAACAAAAAAACCTTATTAACGAAAAATTTACTA	16108
Qy	2892	TCAAAGTGCAAAAAAGGCGAAAAATATCCCGTTAACRAGTTTCATTTAATFAGTATAAC	2951
Dd	16107	TATCCCTATTAAAAATCAAAACTCTCTCATATATTCCTTCAACATAACTATAAAATTA	16048
Qy	2952	GCTATAAAACATTTGCACTTTTATFAAAA	2979
Dd	16047	CAAAAATATTATCCCAATTTTATAACA	16020

RESULT 15

```

> US-10-264-213-27/C
> Sequence 27, Application US/10264213
> GENERAL INFORMATION:
> APPLICANT: Glenn, Matthew
> APPLICANT: Havukkala, Ilkka J
> APPLICANT: Lubbers, Mark William
> APPLICANT: Dekker, James
> TITLE OF INVENTION: Polynucleotides, materials incorporating
> TITLE OF INVENTION: them, and methods for using them.
> FILE REFERENCE: 11000.1043c3
> CURRENT APPLICATION NUMBER: US/10/264, 213
> CURRENT FILING DATE: 2002-10-03
> NUMBER OF SEQ ID NOS: 253
> SOFTWARE: FastSeq for Windows Version 4.0
> SEQ ID NO 27
> LENGTH: 13825
> TYPE: DNA
> ORGANISM: Lactobacillus rhamnosus

```


BEST AVAILABLE COPY

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2003, 21:14:45 ; Search time 3682 Seconds

(without alignments)

13103.301 Million cell updates/sec

Title: US-09-842-484A-1

Perfect score: 2979

Sequence: 1 ttataaactgattaaagaag.....acatttgctttattataaa 2979

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hic:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93.4	3.1	1101	17	CNS00EVL
c 2	86.8	2.9	910	13	BM415636
3	82.8	2.8	996	17	CNS00FUFH
c 4	81.8	2.7	641	17	AQ946120
5	80	2.7	562	13	BI815821
6	78.8	2.6	1092	17	CNS020K7

RESULT 1

CNS00EVL

LOCUS

DEFINITION

Drosophila melanogaster genome survey sequence T7 end of BAC: BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION

AL069706

VERSION

AL069706.1

KEYWORDS

GSS.

SOURCE

Drosophila melanogaster.

ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

1 (bases 1 to 1101)

AUTHORS

Genoscope.

TITLE

Direct Submission

JOURNAL

Submitted (02-JUN-1999)

Genoscope

BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)

COMMENT

- Web : www.genoscope.cns.fr

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial

ALIGNMENTS

CNS00EVL 1101 bp DNA linear GSS 04-JUN-1999

Drosophila melanogaster genome survey sequence T7 end of BAC:

BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

AL069706 GI:4949849

GSS.

Drosophila melanogaster.

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 1101)

Genoscope.

Direct Submission

Submitted (02-JUN-1999)

Genoscope

BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial

AL069706 Drosophil

AL071133 Drosophil

AL402900 T3 end of

AL298972 Tetraodon

EG850019 1024027H0

AG058583 Pan trogl

BH31984 AG-ND-138

AL061936 Drosophil

AL105023 Drosophil

AL063632 Drosophil

AL249932 Tetraodon

AL063921 Drosophil

AL069440 Drosophil

AL514901 AL514901

AL106396 Drosophil

AQ897537 HS_3153-A

AL109126 Drosophil

BI1102 F19C22-T7 I

AL436389 T7 end of

AL285149 Tetraodon

AL053444 Drosophil

AZ682338 ENTKR07TF

AL226115 Tetraodon

AL109318 Drosophil

AL063921 Drosophil

AY068550 Schmidtea

AL069440 Drosophil

AL566565 AL566565

BI0881 F24H6-Sp6.1

AL212733 Tetraodon

BQ876453 AGENCOCURT

AL419462 T3 end of

BG854775 1024040G1

BF272396 GA_EB001

AZ539768 ENTDC17TR

AL106171 Drosophil

AL228940 Tetraodon

AL071865 Drosophil

BM416130 OP21217 M

D	b		659	GTTTAAATMAACATAAAGTAATAAWAAAAAAAATAWAAAAAAAAAAAAAAAAAAAAWAATAAA	718
Qy		2378	ACAAGACTGCTAACGTATTAAAGCACTGGCGCACCATTTAAAGTGAATATTAATAAATTAAGTC	2437	
D	b		719	AAAAAATAWWAAAAATAWAAAAATAAAWTAAATTWATMAAAAAAAAAAAAAAAAAAAWATTT	778
Qy		2438	AGTTAAATCCTAAATGCGAATACATCATTTTTTGATATCATGCACAGCCTATTCGTTAAAA	2497	
D	b		779	RWTTATATWAAWAAWTAAAMWTTTKMWWATWAAATWAAAAAAWAAAAAHWATTAATAAA	838
Qy		2498	ATGACAGCTGCTTATATGAAAAAATATGATGCGCATGAATTTCTCAGCATTAACAC	2557	
D	b		839	AAAAAANAANWNTAWNMTWNAAYTYAAATYAAATWNAWAAAAAATWWWTTTTTTT	898
Qy		2558	ATGATGCGATCGAGAANAATCAATGCCGATCCACCATTAAAAAGCTGATAAAAACCTATT	2617	
D	b		899	TTTTTYYCCSSSCSSCSSCSVCVAAVAATAATWAAATKTWATWTTWATTTT	958
Qy		2618	TAAATGACATGACTTAAGAAGTAGAA	2645	
D	b		959	KAAATWAAWAAAAWATTTTTTTTATWAA	986
RESULT 4					
AQ946120/c					
LOCUS					
DEFINITION					
Sheared DNA-46J23.TR Sheared DNA Trypanosoma brucei linear GSS 27-JAN-2000					
Accession Number					
AQ946120					
Version					
Keywords					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Suh,E., Malek,J., Fujii,C., Gerrard,C., Leech,V., de Jong,P.P., Ullu.E., Melville.S., Donelson.J.					
Fraser.C. and Adams.M.					
Determination of clone end sequences from Trypanosoma brucei GUTat					
TITLE					
10.1 sheared DNA library					
JOURNAL			Unpublished (1999)		
COMMENT			Other_GSSs:	Sheared DNA-46J23.TF	Contact: Najib M. El-Sayed Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: nelsayed@tigr.org
Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared DNA library constructed at TIGR. Clones will be available for distribution through ATCC. Sheared DNA end sequences search page: http://www.tigr.org/tldb/mdb/mdb/Cbdb/ . Seq primer: M13-Reverse Class: shotgun. Location/Qualifiers					
FEATURES					
source					
1..641					
/organism=Trypanosoma brucei"					
/strain="TREU927/4 GUTat 10.1"					
/db_xref="taxon:5691"					
/clone="Sheared DNA-46J23"					
/Clone_lib="Sheared DNA"					
/notes=Vector: pUC19; Site_1: SmaI; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (approx 2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome Shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaubin and B. Barrel, Oxford University Press).					


```
Db 461 AAATATATTGATAAGAAAAATTTTAAATRAAATTTTACTATATAAATTAATCAAAACAAA 520
Qy 2503 AGCTATGCTTATATGAARAAATAT 2526
Db 521 AAATATAATATTATAGTAATAAT 544

RESULT 6
CNS020K7 1092 bp DNA linear GSS 12-MAY-2000
LOCUS Tetraodon nigroviridis genome survey sequence T7 end of clone
DEFINITION 222L1 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL175696
VERSION AL175696.1 GI:7813753
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1092)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
JOURNAL
REFERENCE 2 (bases 1 to 1092)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Unpublished
JOURNAL
REFERENCE 3 (bases 1 to 1092)
AUTHORS Genoscope.
Direct Submission
Submitted (12-APR-2000)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
source
Location/Qualifiers
1..1092
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="222L1"
/clone_lib="G"
/note="Genoscope sequence ID : COAG222CF06LP1-end : T7"

BASE COUNT 383 a 169 c 165 g 262 t 113 others
ORIGIN
Query Match 2.6%; Score 78.8; DB 17; Length 1092;
Best Local Similarity 38.6%; Pred. No. 3e-05;
Matches 179; Conservative 58; Mismatches 227; Indels 0; Gaps 0;

Qy 2189 TTTTCTATCCCAATACATTAACGGCTTAGTGAAACAACTAAACAATATTATGAAATA 2248
Db 621 WTTTAAWWAAAAAATAAAAWTTTWTAAWAAAAAATAAAWTTTAAWAAAAA 680

Qy 2249 ATAAAAATATATTCGTTATTATTACATGTTGATAGACATCTTACACACAGACATCA 2308
Db 681 AAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAWAAAAA 740

Qy 2309 AAAAAGAAATATTCGCTTCTCATATAGCACCAAGTGAATTTTACTAAATATGACA 2368
Db 741 WAAWAAWAAWTTTWTWTAAWAAAAAATAAAWTTTAAWTTTAAWAAW 800

Qy 2369 TCTCATATTACACAGTAATAGACTAATAAAACATGAGGCACATTTAAGTAATTAATA 2428
Db 801 TTWTATTTAAWAAAAANAWTTTAAWTTAAWAAAAAATAAATTTTAAWAAWTTT 860

Qy 2429 AATTAAGTCAGTTAAATCTAAATTTGTAATACATCATTTTGTGTAATCATGACGCCAT 2488
Db 861 AWAATTTWAAATWTAAAAAATAAAATTAATAAAWAAWTTTWTWTWTAAWTTWTTAAA 920

Qy 2489 TCGTTAAAAATGACAGCTATGCTTATATGAAAAAATATGATGTCGGCATGAATTTCTCAG 2548
Db 921 TWTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAWAAAAAATAAAWAAWTTTAAAT 980

Qy 2549 CATTAACACATGATGGATCGAGAAATCAATGCCATCCACCATTAAAAAGCTGATTA 2608
Db 981 TATTTTAAAAAATAAAATATAAATATAAAWAAATAAAAATATTTATATATTAAAAAA 1040

Qy 2609 AAACCTATTTTAAATGACAAATGACTTAAGAAGTATGATGAATGTGAAA 2652
Db 1041 AAAAAAATAAATTAATAAAWTTAAATAAAAAAATAAAAAA 1084

RESULT 7
CNS00EVL/c 1101 bp DNA linear GSS 04-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC;
DEFINITION BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL069706
VERSION AL069706.1 GI:4949849
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osogawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
source
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR29B23"
/clone_lib="RPCI-98"
/note="end : T7"

BASE COUNT 419 a 91 c 60 g 299 t 232 others
ORIGIN
Query Match 2.6%; Score 77.8; DB 17; Length 1101;
Best Local Similarity 33.9%; Pred. No. 4.6e-05;
Matches 220; Conservative 115; Mismatches 308; Indels 6; Gaps 2;

Qy 1886 TGTTCCTTAAACTCAGTGAAGTTGAAAAATTAACATCTTAAATAAAATCTGCTATAACC 1945
Db 1089 TWTTTTWTATTTATATATAATAATTAATTTTWTATATATTAATTAATATATTATT 1030

Qy 1946 CGGTATTACATGGTGATTAACATCCATTAAGAACTCGGCATCAAAAGAAAAACCACTT 2005
Db 1946 TTTTAAWAAAAAATAAATTTTAAWTTAAWAAAAAATAAATTTTAAWAAWTTT 860
```


Db	138	AKKAAKAKKACAASSAKAMKASAKAKASAKKCGKGAATAKGAAKAACAACKCKAAA	197
Qy	1724	CGGATGGTAGCTTAATCGCTAATGTTTACAATTTGGCCGAAATTTTCACGAGAAAACTCA	1783
Db	198	AKAKGASAKKAKSAWAKAKAASAKSASAKSKGMAAKCCACACAACAKAAAAAKAA	257
Qy	1784	CAACGGCTATGATGCTCACCATTTTACATGTTTACGATTTAGAGCTTGGCATTTAAACGG	1843
Db	258	KKSACGAGCKCCATCGACGAAATGAACAACAACAACANNNNNNNNNANVANSAAVNN	317
Qy	1844	ATCGATTTAACGAAATATTTGAAACGGCGCTGGATTTACATGTTTCCTTTAAACCTCAGTG	1903
Db	318	VTTBGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	377
Qy	1904	AGTTTGAAAAATTTAAACATCTTAATAAATTCGTATTAACCGGTATTTACATGGTGATA	1963
Db	378	ANAANAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	437
Qy	1964	ACATATCATTAAGAATCGCGATTCACAAAGAAAAACCATTTTGTGTAGTCAATCAGT	2023
Db	438	ATTANATAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	497
Qy	2024	CATTAATAACAGAGCATCAATTTATATAATTTATGACAAATTTGATGATTTAGATGAAA	2083
Db	498	NATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	557
Qy	2084	GTAGAAAGTATATCTTCAATAAACCCTGCAATATCAAGAGAAAT-----GGATATGT	2137
Db	558	ATTATATTAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	617
Qy	2138	TAAAGATCTTAACTCATTCACAAATAAGATGCGCAAAATCGCAGTCAGTATTTCTATC	2197
Db	618	TATAANAWATAATATATATATATATATATATATATATATATATATATATATATATAT	677
Qy	2198	CCATACATTAACGGCTTAGTGAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	2257
Db	678	TAWATTTTAAANNAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	737
Qy	2258	TATTCGTTTATTTCTACATGTTGATAGAAATCATCTTACACAGACATCAAAAAAGAAA	2317
Db	738	AA-----AAAWTTTTATATATANAANAANAATAAATAAATAAATAAATAAATAAATAA	792
Qy	2318	TATTCGTTTCTATCATAGCACCAAGTGAATTTTACTTAATAATGACATCTCATATT	2377
Db	793	TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	852
Qy	2378	ACACGATTAAGACTAATAAATAACGAGGCACATTTAAGTAATATTAATAAATAAAGTC	2437
Db	853	TTAATAAANAATNTTTATTAANAANAANAATAAATAAATAAATAAATAAATAAATAA	912
Qy	2438	AGTTAAATCTAAATGTGAATACATCATTTTGTGATAATCATGACAGCCGCTATTCGTTAAA	2497
Db	913	TTTGTAANAATAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	972
Qy	2498	A 2498	
Db	973	A 973	
RESULT 11			
BG850019/c			
LOCUS			
DEFINITION			
Chlamydomonas reinhardtii CC-1690, normalized, Lambda Zap II			
ACCESSION			
BG850019			
VERSION			
BG850019.1 GI:14231203			
KEYWORDS			
EST.			
SOURCE			
Chlamydomonas reinhardtii.			
ORGANISM			
Chlamydomonas reinhardtii.			
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;			
Chlamydomonadaceae; Chlamydomonas.			
REFERENCE			
1 (bases 1 to 973)			
AUTHORS			
Grossman, A., Davies, J., Federspiel, N., Harris, E., Lefebvre, P.,			

TITLE			
JOURNAL			
COMMENT			
McDermott, J.P., Silflow, C., Stern, D. and Surzycki, R. Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants: project phase 2 Unpublished (2000) Contact: Charles Hauser DCMB Box 91000 Durham, NC 27708-1000 Tel: 919 613 8159 Fax: 919 613 8177 Email: chauser@duke.edu. Location/Qualifiers 1. 973 /organism="Chlamydomonas reinhardtii" /strain="CC-1690 wild type mt+ 21gr" /db_xref="taxon:3055" /clone_lib="C. reinhardtii CC-1690, normalized, Lambda zap II" /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2: XhoI; This library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (acetate-containing) medium in the light, TAP medium in the dark, HS (minimal) medium in ambient levels of CO2 and HS medium bubbled with 5% CO2. PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into Lambda ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the Lambda ZAP clones by superinfection with EXassist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."			
BASE COUNT			
ORIGIN			
Query Match 2.6%; Score 76.4; DB 12; Length 973; Best Local Similarity 47.3%; Pred. No. 8.7e-05; Matches 262; Conservativity 0; Mismatches 291; Indels 1; Gaps 1;			
Qy	1972	ATTAAGAACTCGGCATTCACAAAGAAACCATTTTCTGTGTAGTCAATCAGTCATTAAT	2031
Db	838	ATTAAGAAATATATAAATAAGATAATTAATTAATAAATAAATAAATAAATAAATAA	779
Qy	2032	AGACAAGGCATCAATTTATTAATTTATGACAAATTTGATGATTTAGATGAAGTAGAAG	2091
Db	778	ATAAATAATTTTATTAATTAATAAATAAATAAATAAAGAGATATAAATAAATAAATAA	719
Qy	2092	TATATCTTCAATAAACCCTGATATCAAGAAGAAATGGATATGTTAAAGATCTTAAA	2151
Db	718	AAATAAATAAAGTATCCAAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAA	659
Qy	2152	CTCATTCAAAATAAAGATGCGCAAAATCGCAGTCAGTATTTTCTATCCCAATACATTAAC	2211
Db	658	ATAATTAATAAATAAAGCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	599
Qy	2212	GGCTTAGTGAAAACTAAACAATTTATGTAATATAATAAATAAATAAATAAATAAATAA	2271
Db	598	GAATTAATAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	539
Qy	2272	CTACATGTTGATAGATCATCTTTACACAGACATCAAAAAATAAATAAATAAATAAATAA	2331
Db	538	AACAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	479
Qy	2332	CATAAGCACCAGTGAATATTTTACTAATAATGACATCTCATATTTACGAGTAAATAGA	2391
Db	478	TA-AATTACAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	420
Qy	2392	CTAATAAATACTGAGGCACATTTAAGTATATTAATAAATAAATAAATAAATAAATAAATAA	2451
Db	419	AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	360
Qy	2452	TGTGATACATCATTTTGTGATATCATGACGCCTTATTCGTTAAAAATGACAGCTATGCT	2511

Qy	2215	TTAGTGAAGAAACATAACAATATTATTGCAATATATAAATAATATATTTCGTATTATTCTTA	2274
Dd	329	AAATTAAAAATAAAAATAAAAAATAAAAAATAAAATTTAAAAAATGTTAAAGAATATA	388
Qy	2275	CATGTTGATAGAAATCATCTTACACCAGACATCAAAAAGAAATATTGGCTTTCTATCAT	2334
Dd	389	AATAAATAAAAAATAAAAAATAAAAAATAAAAAATAAATAAAAAATAAANAAG	448
Qy	2335	AAGCACCACCGTAATATTTTACTAAATATGCACATCTCATATTTACACGAGTAAATAGACTA	2394
Dd	449	AAATAAAGAAGAAATAAAAAATAAATAAATAAATAAATAAATAAAGAGGATAAATATAT	508
Qy	2395	ATAAAACTGAGGCACATTTTAAAGTAATTAATAAATAAAG---TCAGTTAAATCTAAAT	2451
Dd	509	AAATATATAAAAAATGAAATGAATAAANAATAAATAAATAAATAAAGATAAATAAGATATAGAT	568
Qy	2452	TGTCAAATACATCAATTTTGATTAATCATCACAGCCATTCGTTAAAAA	2498
Dd	569	TGTAATAAAATCACATATAATAATAAATAAAGAAGAAATAAATAAANA	615
RESULT	13		
LOCUS	BH391984		
DEFINITION	AG-ND-138G17.TR ND-TAM Anopheles gambiae genomic clone AG-ND-138G17	832 bp	DNA linear GSS 11-DEC-2001
ACCESSION	BH391984		
VERSION	BH391984.1	GI:17338125	
KEYWORDS	GSS.		
SOURCE	African malaria mosquito.		
ORGANISM	Anopheles gambiae		
REFERENCE	1 (bases 1 to 832)		
AUTHORS	Shetty,J., Malek,J., Koo,H., Collins,F., Gardner,M. and Loftus,B.J.		
TITLE	Direct Submission of BAC-end sequences from Anopheles gambiae		
JOURNAL	Unpublished (2001)		
COMMENT	Other.GSSs: AG-ND-138G17.TF.1 Contact: Brendan J Loftus Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0208 Fax: 301 838 3543 Email: bjloftus@tigr.org This clone is from an A. gambiae BAC library (ND-TAM) provided by F.H. Collins and sequenced by The Institute for Genomic Research (TIGR). The BAC library was generated from A. gambiae PEST strain DNA. All DNA was extracted from newly hatched first instar larvae to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M University BAC Center University, College Station, Texas 77843-2123, USA using a HindIII partial digest. Seq primer: M13 Rev Class: BAC ends.		
FEATURES			
source	location/Qualifiers		
	1..832		
	/organism="Anopheles gambiae"		
	/strain="PEST"		
	/db_xref="taxon:7165"		
	/clone="AG-ND-138G17"		
	/clone_lib="ND-TAM"		
	/note="vector: pECBACL; Site_1: HindIII"		
BASE COUNT	425 a 71 c 59 g 277 t		
ORIGIN			
Query Match	2.6%	Score 76;	DB 17; Length 832;
Best Local Similarity	45.1%;	Pred. No. 0.00011;	
Matches 294;	Conservative 0;	Mismatches 340;	Indels 4; Gaps 1;
Qv	1915	TTTAAACATCTTATATAAAATCTGCTATAACCCGCTATTACATGTGTGATAACACATCCATT	1974

Db 150 TTTATCATTCGGTTAAATCTACTAATAATAATAATCAAAATAATAATAAAACAAATAATA 209
QY 1975 AAGAACTCGCATTCACAAAGAAACCACTTTTGTGTAGTCAATCACTGCTTAAATAGA 2034
Db 210 ATAATAACATCATCATATAATAATAATCCACATAATAGCTATATAACAGITATAATAAG 269
QY 2035 CAAGGCATCAATATTATATAATTTATGACAAATTTGTAGTATTTAGATGAAGATAGAAAGTAT 2094
Db 270 AATAATAACATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 329
QY 2095 ATCTTCAATAAACCCTGATATATCAAGAGAAATGGATATGTTTAAAGATCTTAAATC 2154
Db 330 ATTATAAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 389
QY 2155 ATTCAAAATAAAGATGCCAAATCGCAGTCAGTATTTTCTATCCCAATACATTAACCGC 2214
Db 390 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 449
QY 2215 TTAGTGAATAAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 2274
Db 450 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 509
QY 2275 CATGTTGATAAGATCAT---CTTACACGAGACATCAAAAGAAATATTTGGCTTTCTA 2330
Db 510 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 569
QY 2331 TCATAAGCCAAAGTGAATATTTTACTAATAATGACATCTCATATACACGAGTAAAG 2390
Db 570 AAAAAAGATAATTATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 629
QY 2391 ACTAATAAATAACTGAGGCACATTTAAGTAAATAATAATAATAATAATAATAATAATAATA 2450
Db 630 TAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 689
QY 2451 TTGTGATACATCATTTTGTGATAATCATGACAGCCATTTGTTTAAATAAGACAGCTATGC 2510
Db 690 TAATATTAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 749
QY 2511 TTATATGAAAAATAATGATCGCGCATGTAATTTCTCAG 2548
Db 750 ATTATTATGGATATGATATATAATAACATAATCTG 787

RESULT 14
CNS0021J 1101 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR05N11 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL061936
VERSION AL061936.1 GI:4940214
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
TITLE Drosophila melanogaster genome survey sequence TET3 end of BAC #
JOURNAL BP 91 91006 EVRY cedex - FRANCE (E-mail : segre@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org>
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammos in Pletier de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial

ECORI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR05N11"
/clone_lib="RPCI-98"
/note="end : TET3"
/note="*end : TET3"

BASE COUNT 631 a 7 c 28 g 289 t 146 others
ORIGIN

Query Match 2.6%; Score 76; DB 17; Length 1101;
Best Local Similarity 40.1%; Pred. No. 0.0001;
Matches 300; Conservative 40; Mismatches 405; Indels 4; Gaps 1;

QY 1898 TCAGTGAAGTTCGAAATTTTAAACATCTTTAATAAATCTGCTATACCGCGTATTACATG 1957
Db 145 TCATGAAAGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 204
QY 1958 GTGATAACACATCCATTAAGAACTCGGCATTCAGAAAGAAACCACTTTTGTGTAGTCA 2017
Db 205 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 264
QY 2018 ATCAGTCATTAAATAGACAGGATCAATTTATTAATATGACAAATTTGATGATTTAG 2077
Db 265 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 324
QY 2078 ATGAAAGTGAAGTATATCTTCAATAAACCCTGAATATCAAGAGAAATGGATATGT 2137
Db 325 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 384
QY 2138 TAAAGATCTTAACTCATTCATAAATAAAGATGCCAAATCGCAGTCAGTATTTTCTATC 2197
Db 385 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 444
QY 2198 CCAATACATTTAAACGGCTTAGTGAAAACTAAACATATTTTGAATATAATAAATAA 2257
Db 445 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 504
QY 2258 TATTCGTTATATCTACATGTTGATAAGATATCTTACACAGACATCAAAATAAGAAA 2317
Db 505 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 564
QY 2318 TATTCGCTTCTATCATAAGCACCAAGTGAATATTTTACTAAAT---AATGACATCTCA 2373
Db 565 TTTAATTTTAAAWAAATTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 624
QY 2374 TATTACAGGATTAATAGACATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2433
Db 625 TTTTAAAWAATTTTATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 684
QY 2434 AGTCAGTTTAAATCTAAATTTGTAATATACATATTTTGTGATAATCATGACAGCTATCT 2493
Db 685 AATWAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 744
QY 2494 AAAAAATGACAGCTATGCTTATATGAAAAAATAATGATGCGGATCAATTTCTCAGCATTA 2553
Db 745 WATWWTAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 804
QY 2554 ACACATGATTGGATCGAGAAATCAATCGCATCCACCATTTTAAATAAGCTGATTAATAACC 2613
Db 805 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 864
QY 2614 TATTTAATGACATGACTTTAGAAAGTAT 2642
Db 865 ATTTTATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 893

THIS PAGE BLANK (USPTO)

BEST AVAILABLE COPY

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2003, 00:35:34 ; Search time 118 Seconds
(without alignments)
1089.719 Million cell updates/sec

Title: US-09-842-484A-2

Perfect score: 5089

Sequence: 1 MNTLSQAIKAYNSNDYELAL.....SAKKGENVIPNFKFIINSITL 965

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5089	100.0	965	21	AA196213
2	5089	100.0	965	23	AA196213
3	5067	99.6	965	23	AA196213
4	4486.5	88.2	972	20	AA196212
5	4486.5	88.2	972	21	AA196212
6	3079	60.5	702	21	AA196212
7	292.5	5.7	327	23	AA196212
8	276.5	5.4	706	22	AA196212
9	276.5	5.4	715	22	AA196212
10	253	5.0	321	23	AA196212

11	251	4.9	298	22	AA196213	Putative glycosylt
12	246	4.8	322	21	AA196213	CpsII protein whic
13	238.5	4.7	324	18	AA196213	S.thermophilus exo
14	238.5	4.7	324	18	AA196213	S.thermophilus exo
15	230.5	4.5	339	21	AA196213	Enzyme EPS5 involv
16	230.5	4.5	339	21	AA196213	Enzyme EPS5 involv
17	228.5	4.5	322	21	AA196213	Amino acid sequenc
18	228.5	4.5	322	21	AA196213	Enzyme EPS7 which
19	227	4.5	322	21	AA196213	Amino acid sequenc
20	225.5	4.4	316	21	AA196213	Staphylococcus epi
21	225.5	4.4	316	21	AA196213	Enzyme EPS8 which
22	223.5	4.4	278	21	AA196213	Amino acid sequenc
23	223.5	4.4	278	21	AA196213	CpsII protein whic
24	223.5	4.4	674	19	AA196213	Streptococcus pneu
25	217.5	4.3	674	23	AA196213	S. pneumoniae Sp11
26	217.5	4.3	332	21	AA196213	Cps2J protein whic
27	217.5	4.3	334	21	AA196213	Cps2K protein whic
28	217	4.3	330	21	AA196213	Enzyme EPS6 involv
29	216	4.2	336	21	AA196213	Amino acid sequenc
30	216	4.2	336	21	AA196213	Enzyme EPSM involv
31	215.5	4.2	346	23	AA196213	Amino acid sequenc
32	213.5	4.2	327	21	AA196213	Neisseria meningit
33	213.5	4.2	327	21	AA196213	Enzyme EPSJ involv
34	210	4.1	348	17	AA196213	Amino acid sequenc
35	210	4.1	348	18	AA196213	N. gonorrhoeae gly
36	209	4.1	337	17	AA196213	Neisseria polysglyc
37	209	4.1	337	18	AA196213	N. gonorrhoeae gly
38	208.5	4.1	270	21	AA196213	Lipo-oligosacchari
39	207	4.1	301	21	AA196213	Campylobacter jeju
40	206	4.0	328	21	AA196213	Campylobacter jeju
41	206	4.0	329	22	AA196213	Streptococcus pneu
42	205	4.0	389	21	AA196213	EpsM. Lactococcus
43	203	4.0	332	22	AA196213	Campylobacter jeju
44	202	4.0	302	21	AA196213	EpsN. Lactococcus
45	201.5	4.0	324	23	AA196213	Campylobacter jeju
						Lactococcus lactis

ALIGNMENTS

RESULT 1

AA196213

ID AA196213 standard; Protein; 965 AA.

XX AA196213;

AC AA196213;

XX 17-AUG-2000 (first entry)

DE P. multocida chondroitin synthase.

XX Chondroitin synthase; CS: enzyme; hyaluronic acid; ulcer;

XX tissue abrasion; viscoelastic replacement; bioadhesive.

XX Pasteurella multocida.

OS Pasteurella multocida.

XX WO200027437-A2.

XX 18-MAY-2000.

XX 10-NOV-1999; 99WO-US26501.

XX 11-NOV-1998; 98US-0107929.

XX 01-APR-1999; 99US-0283402.

XX (OKLA) UNIV OKLAHOMA STATE.

XX DeAngelis PL;

XX WPI: 2000-376319/32.

XX N-PSDB; AAA27449.

XX Novel method for the enzymatic transfer of sugar molecules to an

XX acceptor, useful for synthesis of e.g. polysaccharide bioadhesives or

PT

drug delivery systems, including hybrid molecules

Claim 15; Page 85; 86pp; English.

The present sequence is the Pasteurella multocida chondroitin synthase Pmcs. Pmcs catalyses glycosaminoglycan polymerisation to produce chondroitin: a linear polysaccharide which has viscoelastic properties which makes it useful for a number of applications. Chondroitin can be used with hyaluronic acid (HA) to coat medical devices e.g. catheters and sensors to reduce tissue abrasion. In addition, they can be used as bioadhesives for haemostatic sealing and healing of wounds and surgical incisions; and as biomaterials that provide sustained delivery of encapsulated drugs, to wounds, ulcers, injuries or surgical sites.

Sequence 965 AA;

Query Match 100.0%; Score 5089; DB 21; Length 965;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 965; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNTLSQAIKAYNSNDYELALKEFSAETVGRKIVFQIICKCKELSTNSYVSEDKKNV 60
DB 1 MNTLSQAIKAYNSNDYELALKEFSAETVGRKIVFQIICKCKELSTNSYVSEDKKNV 60
QY 61 CDSLSLDTATOLLNSVKKLTLSSEKNSLNKWKSTGKSENAEIRKVELVDPKDL 120
DB 61 CDSLSLDTATOLLNSVKKLTLSSEKNSLNKWKSTGKSENAEIRKVELVDPKDL 120
QY 121 VLAPLPHVNDFTWYKRRKSLGKIPVKNIGLSIIPTENRSRIILDTIACLVNOKTNY 180
DB 121 VLAPLPHVNDFTWYKRRKSLGKIPVKNIGLSIIPTENRSRIILDTIACLVNOKTNY 180
QY 181 PFEVVADGSGKENLTIIVQYKOKLDIKYVRQKDYQYQICAVRNGLRTAKYDFVSILD 240
DB 181 PFEVVADGSGKENLTIIVQYKOKLDIKYVRQKDYQYQICAVRNGLRTAKYDFVSILD 240
QY 241 CDMAPOQLVHSHYLTLELNDIVLIGPKRYVDTHNTAEQFLNDPYLESLPETATNN 300
DB 241 CDMAPOQLVHSHYLTLELNDIVLIGPKRYVDTHNTAEQFLNDPYLESLPETATNN 300
QY 301 PSITSKGNISLDWRLEHFKFTDNLRLCDSPPRYVAGNVAFSEKWLKNVGVDFEENHWG 360
DB 301 PSITSKGNISLDWRLEHFKFTDNLRLCDSPPRYVAGNVAFSEKWLKNVGVDFEENHWG 360
QY 361 GEDVEFSYRLFAKCFRVIDGGMAIHQEPGKENETEREAGKSTILKIVKEKVPYIYRK 420
DB 361 GEDVEFSYRLFAKCFRVIDGGMAIHQEPGKENETEREAGKSTILKIVKEKVPYIYRK 420
QY 421 LLPIEDSHIHRPLVSIYIPAYNCANYIQRVCDSALNQTVDLEVCICNDGSTNTLEVI 480
DB 421 LLPIEDSHIHRPLVSIYIPAYNCANYIQRVCDSALNQTVDLEVCICNDGSTNTLEVI 480
QY 481 NKLYGNPNRVRIMSPNGGSIASASNAVSFAKGYVIGOLDDYLEPDAVELCLKEFLKD 540
DB 481 NKLYGNPNRVRIMSPNGGSIASASNAVSFAKGYVIGOLDDYLEPDAVELCLKEFLKD 540
QY 541 KTLACVYITNRNVNPDGSLIANGYNWPEFSREKLTATIAHFRMTTIRAWHLTDGFNEN 600
DB 541 KTLACVYITNRNVNPDGSLIANGYNWPEFSREKLTATIAHFRMTTIRAWHLTDGFNEN 600
QY 601 IENAVDYDMFLKSEVGKFKHLNKCYNRVLHGDNTSIKLGIOKKNHVVNQSLNROG 660
DB 601 IENAVDYDMFLKSEVGKFKHLNKCYNRVLHGDNTSIKLGIOKKNHVVNQSLNROG 660
QY 661 INYNYDKFDDDESRYIFNKTAEYQOEDMDLKDILQNKQDAKIAVSIYPNTLNGLV 720
DB 661 INYNYDKFDDDESRYIFNKTAEYQOEDMDLKDILQNKQDAKIAVSIYPNTLNGLV 720
QY 721 KKLNNIIEYNKNFVILHVDKNHLPDPIKKEILAPYHKHQNILLNNDISYTSNRLIK 780
DB 721 KKLNNIIEYNKNFVILHVDKNHLPDPIKKEILAPYHKHQNILLNNDISYTSNRLIK 780
QY 781 TEAHLNINKLSQLNCEYIIFDNHDSLFVKNDYAYMKKYDVGGMNFSALTHDWEIKIN 840

DB 781 TEAHLNINKLSQLNCEYIIFDNHDSLFVKNDYAYMKKYDVGGMNFSALTHDWEIKIN 840
QY 841 AHPPFKLIKTYFNDNDRSNVKGASOGMFKYALPHELLTIIKEVITSCQSIDSVPEY 900
DB 841 AHPPFKLIKTYFNDNDRSNVKGASOGMFKYALPHELLTIIKEVITSCQSIDSVPEY 900
QY 901 NTEDFWFOFALLILEKKTGHVFNKTSTITYMPWERKLOWNTNQIOSAKKGENIPVKNKFI 960
DB 901 NTEDFWFOFALLILEKKTGHVFNKTSTITYMPWERKLOWNTNQIOSAKKGENIPVKNKFI 960
QY 961 NSITL 965
DB 961 NSITL 965

RESULT 2

AAM47335
ID AAM47335 standard; Protein; 965 AA.

XX XX

AC AAM47335;

DT 22-FEB-2002 (first entry)

XX Pasteurella multocida chondroitin synthase #1.

XX Chondroitin synthase; dermatan sulphate; chondroitin sulphate polymer;
KW eye application; joint application; moisturiser; drug delivery;
KW wound dressing; biocompatible film.

XX Pasteurella multocida.

XX WO200180810-A2.

XX 01-NOV-2001.

XX 25-APR-2001; 2001WO-US13395.

XX 25-APR-2000; 2000US-199538P.

XX (DANG/) DE ANGELIS P L.

XX De Angelis PL;

XX WPI; 2002-049237/06.

XX N-PSDB; ABAU5097.

XX New chondroitin synthase gene obtained from Pasteurella multocida,
DR useful as hyaluronan polysaccharide substitute in medical or cosmetic
XX applications, e.g. for eye or joint applications, for moisturizer or
XX wound dressings

XX Claim 3; Page 119; 125pp; English.

XX The present invention relates to the coding sequence of the Pasteurella
XX multocida chondroitin synthase. A chondroitin polysaccharide may be used
XX as a hyaluronan polysaccharide substitute in medical or cosmetic
XX applications, for example in eye or joint applications, for moisturiser
XX or wound dressings. The enzyme may be used in covalently coupling
XX specific drugs, proteins or toxins to the structurally modified
XX chondroitin for general or targeted drug delivery or radiological
XX procedures, covalently cross linking the hyaluronic acid itself or to
XX other supports to achieve a gel or other three dimensional biomaterial
XX with stronger physical properties, and covalently linking hyaluronic acid
XX to a surface to create a biocompatible film or monolayer. The present
XX sequence is one version of the protein of the invention.

XX Sequence 965 AA;

Query Match 100.0%; Score 5089; DB 23; Length 965;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 965; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGRIKIVFEQIIKCKEKLSTNSYVSEDKKNV 60
Db 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGRIKIVFEQIIKCKEKLSTNSYVSEDKKNV 60
QY 61 CDSLSLDTATOLLNSNVKLLTSLSEKNSLKNKWKSTGKKSSENAEIRKVELVPKDPKDL 120
Db 61 CDSLSLDTATOLLNSNVKLLTSLSEKNSLKNKWKSTGKKSSENAEIRKVELVPKDPKDL 120
QY 121 VLAPLPHVNDFTWYKRNKSLGKIPVKNKNIGLSIIPTFNRSRLDITLACLVNOKTNY 180
Db 121 VLAPLPHVNDFTWYKRNKSLGKIPVKNKNIGLSIIPTFNRSRLDITLACLVNOKTNY 180
QY 181 PFEVVVADGSKENLLTIVQKYEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKYDFVSILD 240
Db 181 PFEVVVADGSKENLLTIVQKYEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKYDFVSILD 240
QY 241 CDMAPOOLWVHSYLTLELLENDIVLIGPRKYVDTHNTAEQFLNDPYLESLPETATNNN 300
Db 241 CDMAPOOLWVHSYLTLELLENDIVLIGPRKYVDTHNTAEQFLNDPYLESLPETATNNN 300
QY 301 PSITSGNISLDRLEHFKKTDNLRLCDSPFYFVAGNVAFSKEWLNKVGWFEDEFNHWG 360
Db 301 PSITSGNISLDRLEHFKKTDNLRLCDSPFYFVAGNVAFSKEWLNKVGWFEDEFNHWG 360
QY 361 GEDVEGYRLFAKCGCFRVIDGMAIHQEPGPKENETEREAGKSITLKIYKVKVPYIYRK 420
Db 361 GEDVEGYRLFAKCGCFRVIDGMAIHQEPGPKENETEREAGKSITLKIYKVKVPYIYRK 420
QY 421 LLPIEDSHIIRIPLVSIYIPAYCANYIQRCDVSALNQTVDLEVCJNDGSDTNTLEVI 480
Db 421 LLPIEDSHIIRIPLVSIYIPAYCANYIQRCDVSALNQTVDLEVCJNDGSDTNTLEVI 480
QY 481 NKLYGNPNRVRIMSKPNGGJASASNAVSFAKGYIYGQLSDDDYLDPAVELCLKEFLKD 540
Db 481 NKLYGNPNRVRIMSKPNGGJASASNAVSFAKGYIYGQLSDDDYLDPAVELCLKEFLKD 540
QY 541 KTLACVYTTNRNVPDGLSIANGYNWPEFSREKLTITAMIAHFRMFTIRAWHLTDGFEN 600
Db 541 KTLACVYTTNRNVPDGLSIANGYNWPEFSREKLTITAMIAHFRMFTIRAWHLTDGFEN 600
QY 601 IENAVDYDFLKLSEVGFKHLNKKICYNRVLHGDNTSIKKLGIOKKNHVYVQSLNRQG 660
Db 601 IENAVDYDFLKLSEVGFKHLNKKICYNRVLHGDNTSIKKLGIOKKNHVYVQSLNRQG 660
QY 661 INYNYDKFDLDESRYIENKTAEYQEEMLKDLKLIQNKDAKIAVSIFYPNTLGLV 720
Db 661 INYNYDKFDLDESRYIENKTAEYQEEMLKDLKLIQNKDAKIAVSIFYPNTLGLV 720
QY 721 KKLNNIIEYNKNIFVILHVDKNHLPDIKKEILAFYHKHQNILLNNDISYYSNRLIK 780
Db 721 KKLNNIIEYNKNIFVILHVDKNHLPDIKKEILAFYHKHQNILLNNDISYYSNRLIK 780
QY 781 TEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSDYAYMKKYDVGMMFNSALTHDWIEKIN 840
Db 781 TEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSDYAYMKKYDVGMMFNSALTHDWIEKIN 840
QY 841 AHPPEFKLIKTFFNDNDRSMNVKASQGMFKYALPHELLTIKEIVITSCQSDSYPEY 900
Db 841 AHPPEFKLIKTFFNDNDRSMNVKASQGMFKYALPHELLTIKEIVITSCQSDSYPEY 900
QY 901 NTEDIWFQFALLILEKKTGHVFNKTSITLYMPWERKIQWTNEQIQSAKKGENTPVNKFII 960
Db 901 NTEDIWFQFALLILEKKTGHVFNKTSITLYMPWERKIQWTNEQIQSAKKGENTPVNKFII 960
QY 961 NSITL 965
Db 961 NSITL 965
```

RESULT 3
AA047336
ID AA047336 standard; Protein: 965 AA.
XX

```
AC AA047336;
XX 22-FEB-2002 (first entry)
XX Pasteurella multocida chondroitin synthase #2.
DE Chondroitin synthase; dermatan sulphate; chondroitin sulphate polymer;
XX eye application; joint application; moisturiser; drug delivery;
KW wound dressing; biocompatible film.
XX Pasteurella multocida.
OS WO200180810-A2.
XX PN 01-NOV-2001.
XX PD 25-APR-2001; 2001WO-US13395.
XX XX 25-APR-2000; 2000US-199538P.
XX (DANG/) DE ANGELIS P L.
XX De Angelis PL;
PI WPI; 2002-049237/08.
XX DR N-PSDB; ABA05098.
XX PT New chondroitin synthase gene obtained from Pasteurella multocida,
PT useful as hyaluronan polysaccharide substitute in medical or cosmetic
PT applications, e.g. for eye or joint applications, for moisturizer or
PT wound dressings.
XX Claim 3; Page 122; 125pp; English.
PS The present invention relates to the coding sequence of the Pasteurella
XX multocida chondroitin synthase. A chondroitin polysaccharide may be used
CC as a hyaluronan polysaccharide substitute in medical or cosmetic
CC applications, for example in eye or joint applications, for moisturiser
CC or wound dressings. The enzyme may be used in covalently coupling
CC specific drugs, proteins or toxins to the structurally modified
CC chondroitin for general or targeted drug delivery or radiological
CC procedures, covalently cross linking the hyaluronic acid itself or to
CC other supports to achieve a gel or other three dimensional biomaterial
CC with stronger physical properties, and covalently linking hyaluronic acid
CC to a surface to create a biocompatible film or monolayer. The present
CC sequence is one version of the protein of the invention.
XX Sequence 965 AA;
```

```
Query Match 99.6%; Score 5067; DB 23; Length 965;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 960; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGRIKIVFEQIIKCKEKLSTNSYVSEDKKNV 60
Db 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGRIKIVFEQIIKCKEKLSTNSYVSEDKKNV 60
QY 61 CDSLSLDTATOLLNSNVKLLTSLSEKNSLKNKWKSTGKKSSENAEIRKVELVPKDPKDL 120
Db 61 CDSLSLDTATOLLNSNVKLLTSLSEKNSLKNKWKSTGKKSSENAEIRKVELVPKDPKDL 120
QY 121 VLAPLPHVNDFTWYKRNKSLGKIPVKNKNIGLSIIPTFNRSRLDITLACLVNOKTNY 180
Db 121 VLAPLPHVNDFTWYKRNKSLGKIPVKNKNIGLSIIPTFNRSRLDITLACLVNOKTNY 180
QY 181 PFEVVVADGSKENLLTIVQKYEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKYDFVSILD 240
Db 181 PFEVVVADGSKENLLTIVQKYEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKYDFVSILD 240
QY 241 CDMAPOOLWVHSYLTLELLENDIVLIGPRKYVDTHNTAEQFLNDPYLESLPETATNNN 300
Db 241 CDMAPOOLWVHSYLTLELLENDIVLIGPRKYVDTHNTAEQFLNDPYLESLPETATNNN 300
```

301 PSITSGNISLDRLEHFKKTDNLRLCDSPRYFVAGNVAFSKEWLNKVGWDFDEEFHWG 360
DB 301 PSITSGNISLDRLEHFKKTDNLRLCDSPRYFVAGNVAFSKEWLNKVGWDFDEEFHWG 360
QY 361 GEDVEFGYRLFAKGCFFRVIDGMAIHQEPGCKENETREAGKSTLTKIVKEKVPYIRK 420
DB 361 GEDVEFGYRLFAKGCFFRVIDGMAIHQEPGCKENETREAGKSTLTKIVKEKVPYIRK 420
QY 421 LPIEDSHIRPLVSIYIPAYNCANYIQCVCDSALNQTVDVLEVCICNDGSTDNTLEVI 480
DB 421 LPIEDSHIRPLVSIYIPAYNCANYIQCVCDSALNQTVDVLEVCICNDGSTDNTLEVI 480
QY 481 NKLYGNPNRVRIMSKPNGGSIASANAASFAKGYIIGQLSDSDYDLEPDAVELCLKEFLKD 540
DB 481 NKLYGNPNRVRIMSKPNGGSIASANAASFAKGYIIGQLSDSDYDLEPDAVELCLKEFLKD 540
QY 541 KTLACVYTTNRNPNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHLTDGFEN 600
DB 541 KTLACVYTTNRNPNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHLTDGFEN 600
QY 601 IENAVDYDMFLKSEYGVKFKHLNKCINRYNLHGDNTSIKKGIOQKNHFFVWNQSLNROG 660
DB 601 IENAVDYDMFLKSEYGVKFKHLNKCINRYNLHGDNTSIKKGIOQKNHFFVWNQSLNROG 660
QY 661 INTYNDKDDLESRYIFNKTAETAEQEEEDMLKDLKIQNKDAKIAVSIFFPNTLNGLV 720
DB 661 INTYNDKDDLESRYIFNKTAETAEQEEEDMLKDLKIQNKDAKIAVSIFFPNTLNGLV 720
QY 721 KKLNNIIEYNKNIFVILHVDKNHLTPDIKKEILAFYKHQVNLNNDISYTSNRLIK 780
DB 721 KKLNNIIEYNKNIFVILHVDKNHLTPDIKKEILAFYKHQVNLNNDISYTSNRLIK 780
QY 781 TEAHLNINKLSQNLNCEYIIFDNHDSLVFNKDSYAYMKKYDVGWNSFALTDHWTIEKIN 840
DB 781 TEAHLNINKLSQNLNCEYIIFDNHDSLVFNKDSYAYMKKYDVGWNSFALTDHWTIEKIN 840
QY 841 AHPEFKLIKTYFNDNDRSMNVKGSQGMFKYALPHELLTIIKEVITSQSIDSVPEY 900
DB 841 AHPEFKLIKTYFNDNDRSMNVKGSQGMFKYALPHELLTIIKEVITSQSIDSVPEY 900
QY 901 NTEDIWFQFALLILEKKTGHVFNKSTLTITMPWERKLOWNEQIOSAKKGENIPVNFKFI 960
DB 901 NTEDIWFQFALLILEKKTGHVFNKSTLTITMPWERKLOWNEQIOSAKKGENIPVNFKFI 960
QY 961 NSITL 965
DB 961 NSITL 965
RESULT 4
ID AAY06212 standard; Protein; 972 AA.
XX AAY06212
AC AAY06212;
XX 16-AUG-1999 (first entry)
DT
DE
XX Pasteurella multocida hyaluronate synthase pmHAS.
XX Hyaluronate synthase; pmHAS; hyaluronic acid; hyaluronan.
XX Pasteurella multocida.
OS
PN W09923227-A2.
PD 14-MAY-1999.
XX 30-OCT-1998; 98WO-US23153.
PF
XX 26-OCT-1998; 98US-0178851.
PR
XX 31-OCT-1997; 97US-0064435.
XX (OKLA) UNIV OKLAHOMA STATE.
PA

DeAngelis P, Kumari K, Weigel PH;
WPI; 1999-337486/28.
N-PSDB; AAX58857.
Nucleic acid encoding hyaluronate synthase for production of
hyaluronic acid with controlled molecular weight and targeting
specificity
Disclosure; Page 123-125; 125pp; English.
This present sequence represents the hyaluronate synthase (pmHAS)
of Pasteurella multocida Carter Type A. The pmHAS enzyme has
different kinetic optima with respect to pH and metal ion
dependence, and different Km values compared with the HAS enzymes
of Streptococcus equisimilis (see AAY06206) and Streptococcus
pyogenes. Km values are about 2- to 3-fold lower for UDP sugars,
and Vmax values are about 2- to 3-fold higher. The invention
provides recombinant vectors containing hyaluronate synthase DNA,
especially S. equisimilis hyaluronate synthase DNA (see AAX58841),
and prokaryotic or eukaryotic host cells which produce the enzyme
and its hyaluronic acid product, particularly a product with
modified structure or molecular size. The hyaluronic acid produced
this way is purer than that produced by conventional methods.
Sequence 972 AA;
Query Match 88.2%; Score 4486.5; DB 20; Length 972;
Best Local Similarity 86.6%; Pred. No. 0;
Matches 842; Conservative 63; Mismatches 60; Indels 7; Gaps 2;
QY 1 MNTLSQAIRKAYNSNDYELALKEKSAETGRKIVFEQIICKKEKL----STNS---YVS 53
DB 1 MNTLSQAIRKAYNSNDYELALKEKSAETGRKIVFEQIICKKEKSAHVSNAHLSVN 60
QY 54 EDKNSVCDSSLDIATQLLSNVKLLTSPSEKNSLNKWKSTGKSENAERKVELVP 113
DB 61 KEEKVNVDSPDLIATQLLSNVKLLTSPSEKNSLNKWKSTGKSENAERKVELVP 120
QY 114 KDFPKDLVLAFLPDHVNDFTYWNRKSLGKIPKVNKNIGLSIIPTNRSRILDTIACL 173
DB 121 KDFPKDLVLAFLPDHVNDFTYWNRKSLGKIPKVNKNIGLSIIPTNRSRILDTIACL 180
QY 174 VNQNTNPFVWVADGSKENLLTIQYEQKLDIKVYKODYQYQLCAVNGLIRTAKY 233
DB 181 VNQNTNPFVWVADGSKENLLTIQYEQKLDIKVYKODYQYQLCAVNGLIRTAKY 240
QY 234 DFVSILDCDAPQOLWVHSYLTLELNDIYVIGPRKYVDTHNTAEOLFNDPYLIESLP 293
DB 241 DFVSILDCDAPQOLWVHSYLTLELNDIYVIGPRKYVDTHNTAEOLFNDPYLIESLP 300
QY 294 ETATNNPSTSKGNISLDWLEHFKKTDNLRLCDSPRYFVAGNVAFSKEWLNKVGWDF 353
DB 301 EVKTNNSVAAKGEFTVSLDWRLEQFEKTNENLRSLDSPRFFFAAGNVAFKWLKSGFFD 360
QY 354 EEFNHGGEDVEFGYRLFAKGCFFRVIDGMAIHQEPGCKENETREAGKSTLTKIVKEK 413
DB 361 EEFNHGGEDVEFGYRLFAKGCFFRVIDGMAIHQEPGCKENETREAGKSTLTKIVKEK 420
QY 414 VPYTYRKLPIEDSHIRPLVSIYIPAYNCANYIQCVCDSALNQTVDVLEVCICNDGST 473
DB 421 VPYTYRKLPIEDSHIRPLVSIYIPAYNCANYIQCVCDSALNQTVDVLEVCICNDGST 480
QY 474 DNTLEVINKLYGNPNRVRIMSKPNGGSIASANAASFAKGYIIGQLSDSDYDLEPDAVELC 533
DB 481 DNTLEVINKLYGNPNRVRIMSKPNGGSIASANAASFAKGYIIGQLSDSDYDLEPDAVELC 540
QY 534 LKEFLKDKTLCACVYTTNRNPNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 593
DB 541 LKEFLKDKTLCACVYTTNRNPNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 600
QY 594 TDGFNENIENAVDYDMFLKSEYGVKFKHLNKCINRYNLHGDNTSIKKGIOQKNHFFVWN 653

Db	841	DWIEKINAHPPKKLIKTYFNDNLKSMNVKGASQGMFTYALAHELLFTIIEVITSCOS	900
QY	894	IDSVPYNTEDWFOALLILEKKTGHVFNKSTLTLYMPWERKLOWTNEQIOSAKKGENI	953
Db	901	IDSVPYNTEDWFOALLILEKKTGHVFNKSTLTLYMPWERKLOWTNEQIOSAKKGENI	960
QY	954	PVNKFIINSITL 965	
Db	961	PVNKFIINSITL 972	
RESULT 6			
ID	AA96212	standard; Protein; 702 AA.	
XX	AA96212;		
AC	AA96212;		
XX	17-AUG-2000	(first entry)	
XX	P. multocida	hyaluronic acid synthase-D.	
XX	Hyaluronic acid synthase-D;	HAS-D; enzyme; ophthalmic surgery;	
KW	cataract; arthritis; ulcer;	tissue abrasion; bioadhesive;	
KW	viscoelastic replacement;	hyaluronic acid production.	
XX	Pasteurella multocida.		
OS			
XX	Key	Location/Qualifiers	
PH	Misc-difference 450		
FT	/note= "encoded by AACGTG"		
XX	WO200027437-A2.		
XX	18-MAY-2000.		
XX	10-NOV-1999;	99WO-US26501.	
XX	11-NOV-1998;	98US-0107929.	
PR	01-APR-1999;	99US-0283402.	
XX	(OKLA)	UNIV OKLAHOMA STATE.	
PA	DeAngelis PL;		
PI	WPI: 2000-376319/32.		
DR	N-PSDB; AA27448.		
XX	Novel method for the enzymatic transfer of sugar molecules to an		
PT	acceptor, useful for synthesis of e.g. polysaccharide bioadhesives or		
PT	drug delivery systems, including hybrid molecules -		
XX	Claim 11;	Page 83; 86pp; English.	
XX	The present sequence is the soluble recombinant form of Pasteurella		
CC	multocida hyaluronic acid synthase-D, PmHAS-D. This sequence encompasses		
CC	residues 1 to 703 of the 972 residues of the native PmHAS. PmHAS-D		
CC	catalyses glycosaminoglycan polymerisation to produce hyaluronic acid,		
CC	HA: a linear polysaccharide. HA has viscoelastic properties which makes		
CC	it useful for a number of applications. HA can be used during ophthalmic		
CC	surgery as a viscoelastic replacement for the vitreous humour e.g. during		
CC	implantation of intraocular lenses in cataract patients. HA injections		
CC	directly into joints is also used to alleviate pain associated with		
CC	arthritis. HA can also be used to coat medical devices e.g. catheters and		
CC	sensors to reduce tissue abrasion. HA can also be used as bioadhesives		
CC	for haemostatic sealing and healing of wounds and surgical incisions; and		
CC	as biomaterials that provide sustained delivery of encapsulated drugs, to		
CC	wounds, ulcers, injuries or surgical sites. The present sequence can		
CC	therefore be used to produce HA.		
XX	Sequence	702 AA;	
SQ	Query Match	60.5%; Score 3079; DB 21; Length 702;	
	Best Local Similarity	82.4%; Pred. No. 1.2e-227;	

```
XX PA (INSP ) INST PASTEUR.
XX PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
XX PI Dussurget O, Chetoui L, Medjari H, Glaser P, Kunst F, Cossart P;
XX PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
XX PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
XX PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
XX PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
XX PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
XX PI Rose M, Voss H;
XX DR WPI; 2002-010914/01.
XX PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
XX PT and prevention of Listeria and related bacterial infections, and
XX PT related polypeptides -
XX PS Claim 6; SEQ ID NO 1270; 192pp; French.
XX CC The present invention relates to the genome sequence of Listeria
XX CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
XX CC it are useful for selecting probes and primers for detecting genes in L.
XX CC monocytogenes and related organisms, and for studying genetic
XX CC polymorphisms and other genomes. The present sequence is a protein
XX CC encoded by the genome sequence of the present invention. Proteins
XX CC expressed from the genome sequence are useful for raising specific
XX CC antibodies, identification of L. monocytogenes and related organisms, and
XX CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
XX CC B12. The genome sequence and proteins encoded by it are also useful for
XX CC selecting compounds that regulate gene expression and cell replication
XX CC and modulate L. monocytogenes-related diseases. In addition, the genome
XX CC sequence and proteins encoded by it are useful in pharmaceutical and
XX CC vaccines compositions for the treatment or prevention of infections by L.
XX CC monocytogenes and related organisms.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 327 AA;
Query Match 5.7%; Score 292.5; DB 23; Length 327;
Best Local Similarity 27.5%; Pred. No. 5.5e-14;
Matches 98; Conservative 44; Mismatches 133; Indels 81; Gaps 9;
Qy 433 PLVSIIYPAYCANYIQCVDYSALNQTVDLEVCICNDGSDTDLTVINKLYGNPNRVRI 492
Db 3 PLVSIIIPVYVNEKYVXRCLDSVLEQTYHNLEVIIVNDGATDNSAKVIKTSIDN--RIRY 60
Qy 493 MSKPNGGTASNAASVFAKGYIQLDSDDDYLEPDAVELCLKEFLKDKTLACVYTTNRN 552
Db 61 FEKENAGQATARNGLDVAATGDYIVMVDSDDYISKNLVETCL-DTVQKTNADLVLTFSYN 119
Qy 553 VNPQGS-----LIANGYNWPEFSREKLTAMIAHHFRMTTIAHHLTDGPF--- 597
Db 120 VNQEGEMQYIKRDKGIKVLDAAGTPWKNFYQADLWKG-----SRFPVGYWYEDLGIIPV 173
Qy 598 -----NENTENAVDY---DMFLKLSVGFKFLWKLKICYNRVLHGDNTSIKKLGIOKN 647
Db 174 VTLKADNPVKIQDALYIITRADSQSIQOQVDFLDVVI--MLENVETELKLGII----- 227
Qy 648 HFVVVNSLNRQGINNYNDKFDLDESRYIFNKTAIYQEMDMKDL---KLIONKDA 704
Db 228 -----YEESKOQLAVLYIEHLIYLRVL 249
Qy 705 KIAYSIFPNTLGLVKKLNNIENKNIFVILLVHDKNHLTPDIIKKELIAPYKH 760
Db 250 RKAIYITNKQKKKLIKIKISTIIQEKFPNWSYPYQAGGKLTATLKKKALWYLHH 305
RESULT 8
AAU33454
ID AAU33454 standard; Protein; 706 AA.
```

```
XX AC AAU33454;
XX DT 14-FEB-2002 (first entry)
XX DE Enterococcus faecalis cellular proliferation protein #90.
XX KW Antisense; prokaryotic cellular proliferation protein;
XX KW antibiotic; antibacterial; drug design.
XX OS Enterococcus faecalis.
XX PN WO200170955-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US09180.
XX PR 21-MAR-2000; 2000US-191078P.
XX PR 23-MAY-2000; 2000US-206848P.
XX PR 26-MAY-2000; 2000US-207727P.
XX PR 23-OCT-2000; 2000US-242578P.
XX PR 27-NOV-2000; 2000US-253625P.
XX PR 16-DEC-2000; 2000US-257931P.
XX PR 16-FEB-2001; 2001US-269308P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX PI Yamamoto RT, Xu HH;
XX DR WPI; 2001-611495/70.
XX DR N-PSDB; AAS51313.
XX PT New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids -
XX PS Example 3; Seq ID No 4950; 511pp; English.
XX CC The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the
XX CC genes, their use in the discovery of novel antibiotics, the essential
XX CC genes themselves and the encoded proteins. The prokaryotes used are
XX CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX CC pneumoniae, pseudomonas aeruginosa and Enterococcus faecalis. The
XX CC invention is also useful for the identification of potential new targets
XX CC for antibiotic development. The antisense nucleic acids can also be used
XX CC to identify proteins used in proliferation, to express these proteins,
XX CC and to obtain antibodies capable of binding to the expressed proteins.
XX CC The proteins can be used to screen compounds in rational drug discovery
XX CC programmes. The antisense nucleic acid sequence is also useful to screen
XX CC for homologous nucleic acids which are required for cell proliferation in
XX CC a wide variety of organisms. The present sequence represents an
XX CC essential prokaryotic cellular proliferation protein.
XX CC Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 706 AA;
Query Match 5.4%; Score 276.5; DB 22; Length 706;
Best Local Similarity 21.5%; Pred. No. 3.1e-12;
Matches 151; Conservative 115; Mismatches 268; Indels 167; Gaps 25;
Qy 260 DNDIVILGPKRYVD--THNITAEQFLNDPYLESLPETATNNPNPITSKGNISLDMR--- 314
Db 2 NEDIKVIQFSDIYRDKATNNLTITGALDITTKES--PTFTINNQVSAYNQIQRVLRDV 59
Qy 315 -----LEHFKKTDNLRLCDSPPFFYFVAGNVAFSKEWLNK-----VGWF 352
Db 60 NQIYQTEPAIEAGFVVTLEGIKQKKVL-----PPHFQSSAHVVTVDFFLNKKYVIPGTE 114
```


Db 490 AKFEQLPGRFVESIDIPF-----NFSTINRAAKAHGEVLLFLNDTEVIT 538
Qy 813 ND-----SYA-----YMKK 821
Db 539 ENWTLMVSAQQBRIGCVGAKLPPNNTVOHAGVILGLGVAGHGHYGYPHGDLYGFR 598
Qy 822 YDVGMFSAETHDWI-----EKINAHPPFKLIKTYFNDNL 858
Db 599 LAINVNSAVTAACLLMKKADFDAVGGEFAFTVAENDVDL 639
RESULT 10
ID ABP26805 standard; Protein; 321 AA.
XX
AC ABP26805;
XX
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 2786.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus agalactiae.
XX
PN W0200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB04789.
XX
PR 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.
PR 07-MAR-2001; 2001GB-0005640.
XX
XX (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
PA
PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
PI Tettelin H;
XX
XX WPI; 2002-352536/38.
DR N-PSDB; ABN67436.
XX
XX New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
PS Claim 1; Page 3429; 4525pp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
XX Streptococcus proteins.
XX
SQ Sequence 321 AA;

Query Match 5.0%; Score 253; DB 23; Length 321;
Best Local Similarity 25.1%; Pred. No. 5.8e-11;
Matches 94; Conservative 64; Mismatches 142; Indels 74; Gaps 15;
Qy 435 VSIYIPAYCANYTORCVDSALNQTVVDLEVCICNDGSTDNTLEVINKLYGNPNRVRIMS 494
Db 4 VSIIPVYVQSFNCEISVLAQTSYLNLEILVNDGSDNSGDICDYSEIDGRIFVFH 63
Qy 495 KPNGGASANAASFAKGYIGOLDSDYL-EPDAVELCL--KEFLKDKTLACVYTTN 550
Db 64 KNGGLSDARNYGISRATGDIYLLDDSDYLYKEDAIERMVEFSKYSEIVLGC-YVEK 122
Qy 551 R-----NVNPDGSLIANGYNPFESREKLTAMIAHFMETIRAWHLTDGFENENIENAV 605
Db 123 REQHIINIVLEDEMI-----ETISPV-----QAIONIIY 150
Qy 606 DYDMFLKLSYGVKFKHLNKKICYNRVLHGDNTSIIKLGIOKNHFFVVVQSLNRQGINYN 665
Db 151 NYDAYRAIFTVAH-----NKL-YKRELF--STLCYPVGLKHEDEFLTYLAKAKNIIFR 203
Qy 666 YDKFDDLDESRYIFNK----TAEQEEMDMKDLKLIONKDAKIAVSIFYPNTL---N 717
Db 204 YNTY-----AYRIRENSIMTGSYN-----IKRLHVAEALKERYLLEKYPDLVQSER 251
Qy 718 GLVKKLN-NIIEYNKNIEVIIHVDKNHLTPDIKKEILAFYHKHQNILLNNDISYTSN 776
Db 252 ALIKTMEVNLIELYKNFYKEFHT---LKTEYKKTIFDFIKKQRMLLKIKYLLAYCVVH 307
Qy 777 RLKTEAHLNSINK 790
Db 308 FKILNCKRKKKINK 321
RESULT 11
AAB96313
ID AAB96313 standard; Protein; 298 AA.
XX
AC AAB96313;
XX
DT 29-OCT-2001 (first entry)
XX
DE Putative glycosyltransferase, involved in cell wall biogenesis #1.
XX Hyperthermophilic archaeon; hyperthermophilic protein.
XX Pyrococcus abyssi.
XX
XX FR2792651-A1.
XX
XX 27-OCT-2000.
XX
XX 21-APR-1999; 99FR-0005034.
XX
XX 21-APR-1999; 99FR-0005034.
XX (CNRS) CNRS CENT NAT RECH SCI.
XX (IFRE-) IFREMER INST FR RECH EXPL MER.
XX
XX Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;
XX Querellou J, Weissenbach J, Saurin W, Heilig R;
XX WPI; 2001-126236/14.
XX
XX New nucleotide sequences isolated from Pyrococcus abyssi encode
PT proteins useful in industry -
XX
XX Claim 7; Pages 981-982; 1657pp; French.
XX
XX The present invention relates to the genomic sequence of Pyrococcus
CC abyssi (see AAF86431 and AAF41223-7) and P. abyssi proteins. P. abyssi is
CC a hyperthermophilic archaeon, which is isolated from deep-sea
CC hydrothermal vents. The present sequence is one such P. abyssi protein.
XX The proteins of the present invention have various potential industrial

CC uses, since the proteins are stable at very high temperatures, some up to
CC 110 degrees centigrade.

CC Note: This patent is in the same patent family as WO200065062, which
CC contains additional sequences as shown in AAB99132-AAB99143,
CC AAH75903-AAH75920 and AAG66436.

XX Sequence 298 AA;

Query Match 4.9%; Score 251; DB 22; Length 298;

Best Local Similarity 24.3%; Pred. No. 7.3e-11;

Matches 81; Conservative 66; Mismatches 136; Indels 50; Gaps 8;

QY 433 PLVSIIPAYNCANYIQCVDLSALNQIVVDLEVCICNDGSDTNTLEVINKLYGNPRVR- 491

Db 4 PIVSVIIPYINRANLLRRAIASVNLNOKFKDFELIVVDASTDNTPEVESI--EDGRIRY 61

QY 492 IMSKPNGGIASASNAVSAFAGYIGQLSDDDYLEDPDAVELCLKEFLK-DKTLACVYTTN 550

Db 62 IRLKNSGGPIARNIGIKKAGRFIALDDDDWLPRLLEVQVRKFNGLKGFVYVGGF 121

QY 551 RNVNPDGSLIANGYNMPEFSREKLTAMIAHFF-----RMFTTRAWHLTDGFNENIENAV 605

Db 122 YVVSQDGRIL--GKRLPK-HRGDIYSHLLKENFISPTLLIRRECFKAGLFPDRLSSSQ 178

QY 606 DYDFMLKLSEVGFKHLNKCINRVLHGDNWTSIKLGIQKKNHFVVVQNSLNROGINYN 665

Db 179 DWDMLRIARYKFDYVDEIIAYYHVKQISF----- 211

QY 666 YDRFDDLDSEKRYIFNKTAEYQEMDMKDLKLIQNKDAKIYVFPNTLNLGLVKLN- 724

Db 212 -----NMKKYIPGRERLIRKHLDIWKNPKILSIHLSQMLLLLSNNTGKGLKYLTY 263

QY 725 --NTIEYKNIFVILHVDKNHLTPDKKEILA 755

Db 264 SIATAPLENYIMILLKALDSRTEYIKRILS 296

RESULT 12

AAI68974
ID AAY68974 standard; Protein; 322 AA.

XX AC AAY68974;

XX DT 30-MAY-2000 (first entry)

XX DE CpsII protein which has glycosyltransferase activity.

XX KW Capsular gene cluster; serotype 1; polysaccharide biosynthesis;
KW capsular component; antigen; regulation; chain length determination;
KW complement-mediated opsonophagocytosis; serotype-specific detection;
KW antigen; vaccine; Streptococcal disease; CpsII; CpsI; CpsIG;
KW CpsII; CpsII; CpsII; CpsII; glycosyltransferase; CP polymerase.

XX OS Streptococcus suis.

XX PN WO200005378-A2.

XX PD 03-FEB-2000.

XX PF 19-JUL-1999; 99WO-NL00460.

XX PR 22-JUL-1998; 98EP-0202465.

XX PR 22-JUL-1998; 98EP-0202467.

XX PA (DIEN-) STICHTING DIENST LANDBOUWKUNDIG ONDERZOE.

XX PI Smith HE;

XX WPI; 2000-195104/17.

XX N-PSDB; AAZ60930.

XX PT New nucleic acid containing the capsular gene cluster of Streptococcus
XX suis, used for serotype-specific detection and to generate antigens or

PT mutants for vaccination -

XX Disclosure; Fig 4; 144pp; English.

XX The proteins AAY68970-76 are encoded by the capsular gene cluster of
CC Streptococcus suis serotype 1. The genes in this cluster are involved
CC in polysaccharide biosynthesis of capsular components and antigens. The
CC proteins have glycosyltransferase activities (CpsII, CpsI, CpsIG,
CC CpsII, CpsII) and CP polymerase activities (CpsII). The capsule confers
CC bacterium resistance to complement-mediated opsonophagocytosis. The
CC gene cluster is used as a source of probes and primers for
CC serotype-specific detection of S. suis and is also useful for
CC recombinant production of the proteins. The proteins are then useful
CC for producing antigens that can be used in vaccines, for controlling
CC or eradicating a Streptococcal disease, in humans or animals,
CC e.g. against S. suis in pigs.

XX Sequence 322 AA;

Query Match 4.8%; Score 246; DB 21; Length 322;

Best Local Similarity 25.3%; Pred. No. 2e-10;

Matches 92; Conservative 58; Mismatches 120; Indels 94; Gaps 15;

QY 434 LVSIIPAYNCANYIQCVDLSALNQIVVDLEVCICNDGSDTNTLEVINKLYGNPRVRIM 493

Db 4 LISVIVPIYNVQDYLDKINSIINQTYNLEVILVNDGSDTNTLEVINKLYGNPRVRIM 63

QY 494 SKPNPGTASASNAVSAFAGYIGQLSDDDYLEDPDAVELCLKEFLKDKTLACVYTTNRNV 553

Db 64 KKLNGGLADARNGLHATGKYAFVSDDY-----TEVAMFERMHD-----NI 107

QY 554 NPDSGLIA-----NGY-----NWPEFSREKLTAMIAHFFMTIRAWHLTDGF 597

Db 108 TEYNADIAEIDFCLVDENGITKRRNSFNHVTRETVKEFLS----- 150

QY 598 NENIENAVDYDMFLKSEVGFKHLNKCINRVLHG-----NTSIKLGIOKKNH-- 648

Db 151 GSNIE-----NVCKLYSRDIKIDKFIQNNRSIGEDLLFNVLNNVTRVVDREYY 206

QY 649 -FVVVQNSLNROGINYNKDFDLDSEKRYIFNKTAEYQEMDMKDLKLIQNKDAKIA 707

Db 207 NYVIRNSSLLINOKFSINNIDLVRLE---NYPFKLKREFSHYF---DAKVIKEK----- 254

QY 708 VSIFYPNTLNLGLVKLNIIEN--KNIEFVILHVDKNHLT--PDIK-KEILAFYHKKHV 762

Db 255 -----VKCLNKMYSTDCDNEFLFLESYKRIYRPFKAKRYLS--RKHLV 300

QY 763 NILL 766

Db 301 TLXL 304

RESULT 13

AAI14078
ID AAW14078 standard; Protein; 324 AA.

XX AC AAW14078;

XX DT 25-SEP-1997 (first entry)

XX DE S.thermophilus exopolysaccharide biosynthesis enzyme EpsH.

XX KW Exopolysaccharide biosynthesis; eps operon; lactic acid bacterium;
KW epsR; epsR; epsB; epsC; epsD; epsE; epsF; epsG; epsH; epsI; epsJ;
KW epsK; epsL; epsM; epsN; epsO; epsP; epsQ; epsR; epsS; epsT; epsU; epsV; epsW; epsX; epsY; epsZ; CCMC I-1590.

XX OS Streptococcus thermophilus strain Sfi6.

XX PN EP750042-A1.

XX PD 27-DEC-1996.

XX XX 20-JUN-1995; 95EP-0201669.

```
XX 20-JUN-1995; 95EP-0201669.
XX (NEST ) SOC PROD NESTLE SA.
XX Mollet B, Stingele F;
XX WPI; 1997-044836/05.
XX N-PSDB; AAT62328.
XX
XX DNA from lactic acid bacteria - encoding enzymes involved in
XX exo-polysaccharide biosynthesis
XX Claim 7; Pages 33-34; 42pp; French.
XX
XX A chromosomal DNA sequence representing practically the whole eps
XX (exopolysaccharide biosynthesis) operon from Streptococcus
XX thermophilus strain Sf16 (deposited as CNCM I-1590) was isolated.
XX The sequence contained 14 open reading frames, one of which was
XX located on the complementary strand. The present sequence is
XX decoded from the epsH open reading frame. The EpsH protein is
XX probably a glycosyl transferase.
XX Novel exopolysaccharides can be produced by transforming lactic
XX acid bacteria which produce exopolysaccharides with DNA encoding
XX at least one enzyme from the eps operon. The new exopolysaccharides
XX are useful for thickening drinks, liquid desserts, yoghurts, soups,
XX ice-creams, coffee creams, sauces, mayonnaise, etc.
XX
XX Sequence 324 AA;
XX
XX Query Match 4.7%; Score 238.5; DB 18; Length 324;
XX Best Local Similarity 23.9%; Pred. No. 7.6e-10;
XX Matches 90; Conservative 58; Mismatches 139; Indels 89; Gaps 13;
XX
XX QY 434 LVSIYIPAYNCANYIQCVDLSALNQTVDVLEVCICNDGSDTNTLEVINKLYGNPNPRVIM 493
XX LISIVIPYNNVEKYLEKCLQSVQNTYNNFEVILVNDGSTDSSLSICEKFVNQDKRFSVF 65
XX
XX QY 494 SKPNGGIASANAASVFAKGYIGOLDSDDYLEDPAVE-----LCLKKFLKDK 541
XX LISIVIPYNNVEKYLEKCLQSVQNTYNNFEVILVNDGSTDSSLSICEKFVNQDKRFSVF 65
XX
XX QY 542 TLACVYTTNRNVNPDGSLIANGYNNWPFESREKLTMTAMIAHFRMFTIRAWHLTDGFNENI 601
XX LISIVIPYNNVEKYLEKCLQSVQNTYNNFEVILVNDGSTDSSLSICEKFVNQDKRFSVF 65
XX
XX QY 602 ENAVDYDMFLKLESEVGKFKHLNKICYNRVLHGDNTSIKKLGIOQKNHFVVVQNQLRQGI 661
XX LISIVIPYNNVEKYLEKCLQSVQNTYNNFEVILVNDGSTDSSLSICEKFVNQDKRFSVF 65
XX
XX QY 662 NYNYDKFDLDESRRKIFNKTAQYQEMDMKDLKIQNDKAKIAVSIFYPNTLGLVK 721
XX LISIVIPYNNVEKYLEKCLQSVQNTYNNFEVILVNDGSTDSSLSICEKFVNQDKRFSVF 65
XX
XX QY 722 KLNNIIEYNKN-IPV-----IILHVDKNHLPDIIKKEILAFYKHQVNIILLNNDISYTSN 776
XX LISIVIPYNNVEKYLEKCLQSVQNTYNNFEVILVNDGSTDSSLSICEKFVNQDKRFSVF 65
XX
XX QY 777 RLKITE--AHLNINIK 790
XX LISIVIPYNNVEKYLEKCLQSVQNTYNNFEVILVNDGSTDSSLSICEKFVNQDKRFSVF 65
XX
XX QY 29-AUG-1997 (first entry)
XX S.thermophilus exopolysaccharide synthesis operon epsI gene product.
XX
```

```
KW Streptococcus thermophilus; operon; exopolysaccharide; synthesis;
KW open reading frame; thickening; drink; liquid dessert; yoghurt; soup;
KW sauce; ice-cream; coffee cream; mayonnaise; glycosyl transferase.
XX
XX Streptococcus thermophilus.
XX EP750043-A1.
XX 27-DEC-1996.
XX
XX 28-DEC-1995; 95EP-0203663.
XX
XX 20-JUN-1995; 95EP-0201669.
XX (NEST ) SOC PROD NESTLE SA.
XX Mollet B, Stingele F;
XX WPI; 1997-044837/05.
XX N-PSDB; AAT73236.
XX
XX DNA of lactic acid bacteria - encoding enzymes involved in
XX exo-polysaccharide biosynthesis
XX Claim 7; Page 14-28; 46pp; French.
XX
XX This sequence represents the protein encoded by the epsI gene from the
XX Streptococcus thermophilus exopolysaccharide (EPS) synthesis operon.
XX The operon has 13 open reading frames (ORF) designated to a protein
XX encoded by the rfbV ORF of the rfb cluster from Salmonella typhimurium
XX and is probably a glycosyl transferase. The EPS are useful for
XX thickening e.g. drinks, liquid desserts, yoghurts, soups, ice-creams,
XX coffee creams, sauces, mayonnaise.
XX
XX Sequence 324 AA;
XX
XX Query Match 4.7%; Score 238.5; DB 18; Length 324;
XX Best Local Similarity 23.9%; Pred. No. 7.6e-10;
XX Matches 90; Conservative 58; Mismatches 139; Indels 89; Gaps 13;
XX
XX QY 434 LVSIYIPAYNCANYIQCVDLSALNQTVDVLEVCICNDGSDTNTLEVINKLYGNPNPRVIM 493
XX LISIVIPYNNVEKYLEKCLQSVQNTYNNFEVILVNDGSTDSSLSICEKFVNQDKRFSVF 65
XX
XX QY 494 SKPNGGIASANAASVFAKGYIGOLDSDDYLEDPAVE-----LCLKKFLKDK 541
XX LISIVIPYNNVEKYLEKCLQSVQNTYNNFEVILVNDGSTDSSLSICEKFVNQDKRFSVF 65
XX
XX QY 662 NYNYDKFDLDESRRKIFNKTAQYQEMDMKDLKIQNDKAKIAVSIFYPNTLGLVK 721
XX LISIVIPYNNVEKYLEKCLQSVQNTYNNFEVILVNDGSTDSSLSICEKFVNQDKRFSVF 65
XX
XX QY 722 KLNNIIEYNKN-IPV-----IILHVDKNHLPDIIKKEILAFYKHQVNIILLNNDISYTSN 776
XX LISIVIPYNNVEKYLEKCLQSVQNTYNNFEVILVNDGSTDSSLSICEKFVNQDKRFSVF 65
XX
XX QY 777 RLKITE--AHLNINIK 790
XX LISIVIPYNNVEKYLEKCLQSVQNTYNNFEVILVNDGSTDSSLSICEKFVNQDKRFSVF 65
XX
XX QY 29-AUG-1997 (first entry)
XX S.thermophilus exopolysaccharide synthesis operon epsI gene product.
XX
```

RESULT 15
AAW22177
ID AAW22177 standard; Protein; 324 AA.
AC AAW22177;
XX
XX 29-AUG-1997 (first entry)
XX S.thermophilus exopolysaccharide synthesis operon epsI gene product.
XX

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2003, 02:22:39 ; Search time 66 seconds
(without alignments)
430.199 Million cell updates/sec

Title: US-09-842-484A-2

Perfect score: 5089

Sequence: 1 MNTLSQAIKAYNSNDYELAL.....SARKGENIPNFKFIINSITL 965

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCRUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5089	100.0	965	4	US-09-437-277-3
2	3079	60.5	702	4	US-09-437-277-1
3	238.5	4.7	324	1	US-08-597-236-10
4	238.5	4.7	324	1	US-08-746-682A-10
5	227	4.5	727	4	US-09-134-001C-4067
6	223.5	4.4	674	4	US-08-961-083-200
7	210	4.1	348	1	US-08-312-387B-3
8	210	4.1	348	1	US-08-312-387B-11
9	210	4.1	348	1	US-08-683-426-3
10	210	4.1	348	1	US-08-683-426-11
11	210	4.1	348	1	US-08-683-458-3
12	210	4.1	348	1	US-08-683-458-11
13	210	4.1	348	2	US-08-878-360-3
14	210	4.1	348	2	US-08-878-360-11
15	210	4.1	348	3	US-08-478-140B-3
16	210	4.1	348	3	US-08-478-140B-8
17	210	4.1	348	4	US-09-333-412-3
18	210	4.1	348	4	US-09-333-412-11
19	210	4.1	348	4	US-09-338-943-3
20	210	4.1	348	4	US-09-338-943-8
21	209	4.1	337	1	US-08-312-387B-5
22	209	4.1	337	1	US-08-312-387B-12
23	209	4.1	337	1	US-08-683-426-5
24	209	4.1	337	1	US-08-683-426-12
25	209	4.1	337	1	US-08-683-458-5
26	209	4.1	337	1	US-08-683-458-12
27	209	4.1	337	2	US-08-878-360-5

28	209	4.1	337	2	US-08-878-360-12	Sequence 12, Appl
29	209	4.1	337	3	US-08-478-140B-5	Sequence 5, Appl
30	209	4.1	337	4	US-09-333-412-5	Sequence 5, Appl
31	209	4.1	337	4	US-09-333-412-12	Sequence 12, Appl
32	209	4.1	337	4	US-09-338-943-5	Sequence 5, Appl
33	178.5	3.5	990	2	US-08-392-625-20	Sequence 20, Appl
34	178.5	3.5	990	2	US-08-466-961A-20	Sequence 20, Appl
35	176.5	3.5	990	2	US-08-645-193B-15	Sequence 15, Appl
36	174.5	3.4	2710	1	US-08-480-604A-6	Sequence 6, Appl
37	174.5	3.4	2710	2	US-08-405-496A-6	Sequence 6, Appl
38	174.5	3.4	2710	4	US-08-915-136-6	Sequence 6, Appl
39	174.5	3.4	2710	4	US-08-957-310-6	Sequence 6, Appl
40	173.5	3.4	281	4	US-08-961-083-196	Sequence 196, App
41	167.5	3.3	270	4	US-08-961-083-198	Sequence 198, App
42	156.5	3.1	633	2	US-08-648-298-2	Sequence 2, Appl
43	156.5	3.1	10182	4	US-09-134-001C-3159	Sequence 3159, Ap
44	156	3.1	418	4	US-09-134-001C-4051	Sequence 4051, Ap
45	153.5	3.0	956	4	US-09-134-001C-4452	Sequence 4452, Ap

ALIGNMENTS

RESULT 1

US-09-437-277-3

; Sequence 3, Application US/09437277

; Patent No. 644447

; GENERAL INFORMATION:

; APPLICANT: THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA

; TITLE OF INVENTION: POLYMER GRAFTING BY POLYSACCHARIDE SYNTHASES

; FILE REFERENCE: 5820.551

; CURRENT APPLICATION NUMBER: US/09/437,277

; CURRENT FILING DATE: 1999-11-10

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: WordPerfect 8.0 (saved in ASCII format)

; SEQ ID NO 3

; LENGTH: 965

; TYPE: PRT

; ORGANISM: Pasteurella multocida

US-09-437-277-3

Query Match 100.0%; Score 5089; DB 4; Length 965;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 965; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MNTLSQAIKAYNSNDYELALKLFKSAETYGKIVFEQIIRCKEKLSTNSYVSEDKNSV	60
Db	1	MNTLSQAIKAYNSNDYELALKLFKSAETYGKIVFEQIIRCKEKLSTNSYVSEDKNSV	60
Qy	61	CDSSLDIATQLLSNVKLTLSSEKSLKNKWSITCKKSENAIRKVELVPKDFPKDL	120
Db	61	CDSSLDIATQLLSNVKLTLSSEKSLKNKWSITCKKSENAIRKVELVPKDFPKDL	120
Qy	121	VLAPLDHVNDFTWYKNRKSILGKIPVKNIGLSIIIPTFNRSRLDITLACLNVQKTNV	180
Db	121	VLAPLDHVNDFTWYKNRKSILGKIPVKNIGLSIIIPTFNRSRLDITLACLNVQKTNV	180
Qy	181	PFEVWVADGSKENLLITVQKYEOKLDIKYVRQKDYQOLCAVRNLGLRTAKYDFVSILD	240
Db	181	PFEVWVADGSKENLLITVQKYEOKLDIKYVRQKDYQOLCAVRNLGLRTAKYDFVSILD	240
Qy	241	CDMAPOQLVWHSYLTLELLEDNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLPETATNNN	300
Db	241	CDMAPOQLVWHSYLTLELLEDNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLPETATNNN	300
Qy	301	PSITSGKNISLDWRLEHFKKTDNLRCLDSPFRFYFVAGNVAFSKEWLNVKGVDFEENHWG	360
Db	301	PSITSGKNISLDWRLEHFKKTDNLRCLDSPFRFYFVAGNVAFSKEWLNVKGVDFEENHWG	360
Qy	361	GEDVEFGYRLFAKCGFFRVIDGGMAIHOPPEKGENETREAGKSTILKIVKEKVPYIYRK	420
Db	361	GEDVEFGYRLFAKCGFFRVIDGGMAIHOPPEKGENETREAGKSTILKIVKEKVPYIYRK	420

```
QY 421 LLPEDSHIRPLVSIYPAYNCANYIQCVDLSALNQTVDLEVCICNDGSTDNTLEVI 480
|||||
Db 421 LLPEDSHIRPLVSIYPAYNCANYIQCVDLSALNQTVDLEVCICNDGSTDNTLEVI 480
|||||
QY 481 NKLYGNPRVRIMSKPNGGASNAASVAFKGYIIGQSDSDYDLEPDVAVELCKEFLKD 540
|||||
Db 481 NKLYGNPRVRIMSKPNGGASNAASVAFKGYIIGQSDSDYDLEPDVAVELCKEFLKD 540
|||||
QY 541 KTLACVYTTNRNVPDGLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHLTDGFNEN 600
|||||
Db 541 KTLACVYTTNRNVPDGLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHLTDGFNEN 600
|||||
QY 601 IENAVDYDMLKLESEYGVKFKHLNKCYNRYLHGDNTSIKKLGIOKKNHVVVNSLNROG 660
|||||
Db 601 IENAVDYDMLKLESEYGVKFKHLNKCYNRYLHGDNTSIKKLGIOKKNHVVVNSLNROG 660
|||||
QY 661 INYNYDKFDLDESRRYIFNKTAEQEEMDMKDLKLQNKDAKIAVSIFFYPTNLGLV 720
|||||
Db 661 INYNYDKFDLDESRRYIFNKTAEQEEMDMKDLKLQNKDAKIAVSIFFYPTNLGLV 720
|||||
QY 721 KLLNNIEYNKNIFFVILHVDKKNLTPDIKKEILAFYKHQVNNILLNNDISYTSNRLIK 780
|||||
Db 721 KLLNNIEYNKNIFFVILHVDKKNLTPDIKKEILAFYKHQVNNILLNNDISYTSNRLIK 780
|||||
QY 781 TEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDGMNFSALTHTDWTIEKIN 840
|||||
Db 781 TEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDGMNFSALTHTDWTIEKIN 840
|||||
QY 841 AHPFCKLIKTYFNDNDRSMNVKGSQGMFKYALPHELLTIKEVITSCQSDISDYPEY 900
|||||
Db 841 AHPFCKLIKTYFNDNDRSMNVKGSQGMFKYALPHELLTIKEVITSCQSDISDYPEY 900
|||||
QY 901 NTEDIWFQFALLILEKKTGHVFNKTSLTITYPWPKRKLQWNEQIQSAKKGENIPVNFKFI 960
|||||
Db 901 NTEDIWFQFALLILEKKTGHVFNKTSLTITYPWPKRKLQWNEQIQSAKKGENIPVNFKFI 960
|||||
QY 961 NSITL 965
|||||
Db 961 NSITL 965
|||||

RESULT 2
US-09-437-277-1
; Sequence 1, Application US/09437277
; Patent No. 644447
; GENERAL INFORMATION:
; APPLICANT: THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA
; TITLE OF INVENTION: POLYMER GRAFTING BY POLYSACCHARIDE SYNTHASES
; FILE REFERENCE: 5820.551
; CURRENT APPLICATION NUMBER: US/09/437.277
; CURRENT FILING DATE: 1999-11-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: WordPerfect 8.0 (saved in ASCII format)
; SEQ ID NO 1
; LENGTH: 702
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-09-437-277-1

Query Match 60.5%; Score 3079; DB 4; Length 702;
Best Local Similarity 82.4%; Pred. No. 1.1e-260;
Matches 579; Conservative 58; Mismatches 58; Indels 8; Gaps 3;

QY 1 MNTLSQAIKAYNSNDYELAKLFKESAEYIGRKIVFEQIKCKEKL----STNS---YVS 53
|||||
Db 1 MNTLSQAIKAYNSNDYELAKLFKESAEYIGRKIVFEQIKCKEKL---SAHPVSNSAHLN 60
|||||
QY 54 EDKNSVCDSSLDIATQLLSNVKKLTLSSEKNSLKNKWKISITGKKSNAELRKVELYP 113
:::|||||
Db 61 KEEKNVCDSPDLIATQLLSNVKKLTLSSEKNTLKNKWKLTTEKKSNAEVRVALVP 120
|||||
QY 114 KDFPKDLVLAPLPDHVNDFTWYKRRKKSIGIKPVNKNIGLSIITPTFNRSLIDITLACL 173
|||||
```

```
Db 121 KDFPKDLVLAPLPDHVNDFTWYKRRKKSIGIKPEHQHGLSVIITTFNRPAILSLITLACL 180
|||||
QY 174 VNQKNTYPEVVVADGSGKENLTTVQKYEOKLDIKYVRQKDYGYQCALVRNLGLRTAKY 233
|||||
Db 181 VNQKTHYPEVIVTDGSOEDLSPRIQENKLDIRYVRQKNGQFQASARNWGLRLAKY 240
|||||
QY 234 DFEVSIDCDMAQQQWVSHSYLTLELDDNDIVLIGPRKYVDTHNTAEQFLNDPYLESLP 293
|||||
Db 241 DFIGLLDCDMAFNPLWVSHSYVAELLEDLDTIIGPRKIDTQIHDPKDFLNNASLESPL 300
|||||
QY 294 ETATNNNPISITSGNITSLDWRLEHFKKTONLRCLDSPFRFYFAGNVAFSKWLNKVGWFD 353
|||||
Db 301 EVKTNNSVAAKGGTYSLDWRLEQFEKTEENLRSLDSPFRFFAAGNVAFKAKWLNKSGFFD 360
|||||
QY 354 EEFNHGGEDEVGYRLFAKGFVRIDGMAIHQBPFGPKENETEREAKSITLKIIVKEK 413
|||||
Db 361 EEFNHGGEDEVGYRLFRYGSFKTIDGIMAYHQEPFGPKENETDREAGKNITLDMREK 420
|||||
QY 414 VPYIYRKLPIEDSHIRPLVSIYPAYNCANYIQCVDLSALNQTVDLEVCICNDGST 473
|||||
Db 421 VPYIYRKLPIEDSHIRNRPVLSIYPAYN-ANYIQCVDLSALNQTVDLEVCICNDGST 479
|||||
QY 474 DNTLEVINKLYGNPRVRIMSKPNGGASNAASVAFKGYIIGQSDSDYDLEPDVAVELC 533
|||||
Db 480 DNTLEVINKLYGNPRVRIMSKPNGGASNAASVAFKGYIIGQSDSDYDLEPDVAVELC 539
|||||
QY 534 LKFEFLDKTLTACVYTTNRNVPDGLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 593
|||||
Db 540 LKFEFLDKTLTACVYTTNRNVPDGLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 599
|||||
QY 594 TDGFENENIENAVDYDMLKLESEYGVKFKHLNKCYNRYLHGDNTSIKKLGIOKKNHVVVN 653
|||||
Db 600 TDGFENIENAVDYDMLKLESEYGVKFKHLNKCYNRYLHGDNTSIKKLGIOKKNHVVVN 659
|||||
QY 654 QSLNRQGINYNYDKFDLDESRRYIFNKTAEQEEMDMKDL 696
|||||
Db 660 QSLNRQGINYNYDKFDLDESRRYIFNKTAEQEEDILKDI 702
|||||

RESULT 3
US-08-597-236-10
; Sequence 10, Application US/08597236
; Patent No. 5733765
; GENERAL INFORMATION:
; APPLICANT: STINGELE, Francesca
; APPLICANT: MOLLET, Beat
; TITLE OF INVENTION: LACTIC BACTERIA PRODUCING
; TITLE OF INVENTION: EXOPOLYSACCHARIDES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americans
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/597,236
; FILING DATE:
; CLASSIFICATION: 426
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95201669.9
; FILING DATE: 20-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Tanucci A., Allan
; REGISTRATION NUMBER: 30256
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
```


Db 69 IDLDENSGHARNIALERBETPEMFLDADDELASAIITFYLEKFNNTDGLIAPHST-128
QY 549 THRNVPDGLSIANGY-----NWPEFSREKLTAMIAHFRMFTIRAHHLTDGFMENIEN 603
Db 129 TORPOVDLDRVRVEYFNAKENINSFLRKQSACNII---FRTAIRAHHR--FNEMLNT 183
QY 604 AVYDMFLKSLSEGVKFKHLKIKYNRVLHGDNTSIKGLGIQKKNHFVVVYNQSLNRQGINY 663
Db 184 YDWSVLEY-----MKYVNF-----VRI 203
QY 664 YN-----YDKFDDLDSESKRYIFNKTAEOEMDM-L-KDLKLIQNDKAIANSIFY- 712
Db 204 FNPFFFRGEVDPFETLTL-----EQNFDILFKDY-----VNSFYD 241
QY 713 -----PNTLGNLCKLKNII-----EYKNKIFVILHVDKHLNLTDPDKKEILAFYH 758
Db 242 AKRATNPVKREFIVTKMGNKIANEPETRYDIN-----ERYOHTKDLVELSKFLH 293
QY 759 KHOVN-----ILLNNDI-SYITSNRLIKTEAHLNSI-----NKLSQL----- 794
Db 294 VHLVKQKLINKLETILLNNNETDKAFVQFRKTLRHVKNIVLRRKNKERSLYDLTDKE 353
QY 795 -NUNCEYIIFDNDHSLFVNDS-----YAYMKY----- 822
Db 354 DNVKPKTIIVPESFGG---KNYDSPKIYIYEMQKYYPNRYIWSFKNPKNVVPVGSAEKV 410
QY 823 -----DVGMNFSALTHTDWTIEKINAHPPF---KKLIKTYFND-----NDLRSMNVK 864
Db 411 KRNSAEYQAYSEASH-WVS--NARTPLYLNKKENQTYIQTWHTGTPKRLANDKMKVVRMP 467
QY 865 GAS 867
Db 468 GTT 470

RESULT 6
US-08-961-083-200
; Sequence 200, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: Hp Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 200:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 674 amino acids
; TYPE: amino acid

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-961-083-200
Query Match 4.4%; Score 223.5; DB 4; Length 674;
Best Local Similarity 20.1%; Pred. No. 7.8e-11;
Matches 96; Conservative 89; Mismatches 177; Indels 115; Gaps 17;
QY 451 CVDLSALNQTVVLEVCICNDGSDTNTLEVINLKYGNPRVRIMSKPNGGIASASNAVSF 510
Db 1 CLDSITQYKNEIIVVNDGSTDASGEICKFESEMDHRLIYIEQENAGLSAARNTGLNN 60
QY 511 AKGYIYQGLSDSDYLEPDAVELCLK---BFLKDKTKIACVYTTNRN----- 552
Db 61 MSGNYVTFVDSDDWIEQDYVETLYKKIIVEYQADIAVGNYSFNESEGMYFFHILGDSYVE 120
QY 553 -VNPDSLSLANGYNWPEFSREKLTAMIAHFRMFTIRAWHL-----TDGFNENIE 602
Db 121 KVDNVISIFENLYE-----TOEMKSFALISAWGKLYKARLFEQLRFDIGKLGEDGY----- 171
QY 603 NAVDYDMFLKSLSEGVKFKHLNKICY-NRVLHGD-----NTSIKKL 641
Db 172 --LNQKVYL-LSE--KVIVLNKSLYAVIRKSGLSRVWTEKMMHALVDAMSERITILLANM 226
QY 642 GIOKKNHFFVVVNSLRQGINYNYDKFDDLSRKYIFNKTAEOEMDMKDLKLION 701
Db 227 GYPLERHLAVYROMLEVSLAN---GQASGLSDTATY---REFEMKORLLNQLSRQEE 277
QY 702 KDAKIAVISFYPTNLGLVKKLNNIETYNKNIFFVILHVD-KNHLTPDIKKEILAF----- 756
Db 278 SEKKAIVLAANYGYVDQVLTIKSICYHNSIRFYLHSDFPNEWIKQLNKRLEKPDSEI 337
QY 757 ---YHKHQVNILLNNDISYITSNRLIKTEAHLNSINKLSQNLNLCYIIFDNDHSLFVK 812
Db 338 INCRVTSEQIS-CYKSDISYTVELRVFIADF---VQEDKALYLDCLVVTYKLLDLDFAT 392
QY 813 N-DSY-----AYMKKYDVGMNFSALTHTDWTIEKIN 840
Db 393 DLQDYPLAAVYRDFGGRAYFGQEIFNAGVLLVNNAFWKKENNTOKLIDVTNEWHDKVD 449

RESULT 7
US-08-312-387B-3
; Sequence 3, Application US/08312387B
; Patent No. 5545553
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,387B
; FILING DATE: July 7, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800

```
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-312-387B-3

Query Match          4.18; Score 210; DB 1; Length 348;
Best Local Similarity 26.4%; Pred. No. 4.1e-10;
Matches 77; Conservative 57; Mismatches 110; Indels 48; Gaps 14;

QY 433 PLVSIYIPAYNCANYIQRCDVSAALNQTVVDLEVCICNDGSDTNTLEVLKLYGNPNRVRI 492
      ||||| : ||| : | : : : ||| : | : | : ||||| : | : : | : ||| :
Db 3 PLVSVLICAYNVKIFYAGSLAAVYNQTNRLDILVDDGSDGTGLAIKDFKQKRSRIKI 62
      : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 493 MSK-PNGGIASASNAAV-SFAK-----GYIGQLSDSDYLEPDVAVELCLKEFLKDKTLAC 545
      : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 63 LAQAQNSGLIPSLNIGLDELAKSGGGGEYIARTDADDIASPGWIEKIVGEMEKDRSIIA 122
      : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 546 V--YTTNRNVPDGLIA-----NGYNWPEFSREKLTMTMI-----AHFRMTTIRAWHLT 594
      : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 123 MGAWLEVLSEEDGNRLARHHKHKIKWKPTRHEDIAAFFPFGNPINNTMTIMRRS--VI 180
      : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 595 DG---FENENIENAVDYDMFLKLSEVKGPKHLNKICYNRVLHGDNTSIK-----KLGI 643
      || : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 181 DGLLYDTERDWAEDYQFWYDVSKLGRLAYYPEALVKYRLHANQVSSKHSVRQHEIAOGI 240
      : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 644 QK--KNHFVVVYNQSLNRQGINVYNDKFDLDESRYIFNKTAEQYQ-EEMDM 692
      || : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 241 QKTARNDFL---QSMGFK-----TRFDSLE---YRQTKAAAYELPEKDL 278

RESULT 8
US-08-312-387B-11
; Sequence 11, Application US/08312387B
; Patent No. 5545553
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,387B
; FILING DATE: July 7, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-312-387B-11

Query Match          4.18; Score 210; DB 1; Length 348;
Best Local Similarity 26.4%; Pred. No. 4.1e-10;
Matches 77; Conservative 57; Mismatches 110; Indels 48; Gaps 14;

QY 433 PLVSIYIPAYNCANYIQRCDVSAALNQTVVDLEVCICNDGSDTNTLEVLKLYGNPNRVRI 492
      ||||| : ||| : | : : : ||| : | : | : ||||| : | : : | : ||| :
Db 3 PLVSVLICAYNVKIFYAGSLAAVYNQTNRLDILVDDGSDGTGLAIKDFKQKRSRIKI 62
      : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 493 MSK-PNGGIASASNAAV-SFAK-----GYIGQLSDSDYLEPDVAVELCLKEFLKDKTLAC 545
      : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 63 LAQAQNSGLIPSLNIGLDELAKSGGGGEYIARTDADDIASPGWIEKIVGEMEKDRSIIA 122
      : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 546 V--YTTNRNVPDGLIA-----NGYNWPEFSREKLTMTMI-----AHFRMTTIRAWHLT 594
      : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 123 MGAWLEVLSEEDGNRLARHHKHKIKWKPTRHEDIAAFFPFGNPINNTMTIMRRS--VI 180
      : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 595 DG---FENENIENAVDYDMFLKLSEVKGPKHLNKICYNRVLHGDNTSIK-----KLGI 643
      || : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 181 DGLLYDTERDWAEDYQFWYDVSKLGRLAYYPEALVKYRLHANQVSSKHSVRQHEIAOGI 240
      : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 644 QK--KNHFVVVYNQSLNRQGINVYNDKFDLDESRYIFNKTAEQYQ-EEMDM 692
      || : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 241 QKTARNDFL---QSMGFK-----TRFDSLE---YRQTKAAAYELPEKDL 278

RESULT 9
US-08-683-426-3
; Sequence 3, Application US/08683426
; Patent No. 5705367
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,426
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/312,387
; FILING DATE: September 26, 1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-683-426-3
```

Query Match		4.1%; Score 210; DB 1; Length 348;
Best Local Similarity		26.4%; Pred. No. 4.1e-10;
Matches		77; Conservative 57; Mismatches 110; Indels 48; Gaps 14;
QY	433	PLVSIYPAYNCANYIQCVCDSALNOTVVDLEVCICNDGSDTNTLEVINKLYGNNPRVRI 492
DB	3	PLVSVLICAYNVEKYFAQSLAAVNTWRNLDILIVDDGSTDGTGLAIKDFQRDSRIKI 62
QY	493	MSK-PNGGIASASNAV-SFAK-----GYIQLGSDDDYLEPDAVELCLKEFLKDKTLAC 545
DB	63	LAQAQNSGLIPLSINIGLDELAKSGGGGGEYIARTDADDIASPGWIEKIVGEMEKDRSIIA 122
QY	546	V--YTTNRNVNPDGSLIA-----NGYNWPFERSREKLITAMI-----AHHFRMFTIRAWHLT 594
DB	123	MGAWLEVLSEKDGNRRLARHHKKGKIWKPTRHEDIAAFFFFGNPHNNTMIMRRS--VI 180
QY	595	DG---FNENITENAVDYDMFLKSEVGFKFKHLNKICYNRVHLHGDNTSIK-----KLG 643
DB	181	DGGLRYDTERDAEDYQFWDVSKGLRAYYPEALVKYRLHANQVSSKHSVRQHEIAQGI 240
QY	644	QK--KNHFVVVNSLNROGINNYNDKFDLDESRYKIFNKTAEYQ--EEMDM 692
DB	241	QKTARNDFL---QSMGFK-----TRFDSLE-----YRQTKAAAYELPEKDL 278
RESULT 10		
US-08-683-426-11		
; Sequence 11, Application US/08683426		
; Patent No. 5705367		
; GENERAL INFORMATION:		
; APPLICANT: Gotschlich, Emil C.		
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF		
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM		
; NUMBER OF SEQUENCES: 12		
; CORRESPONDENCE ADDRESS:		
; ADDRESSEE: Klauber & Jackson		
; STREET: 411 Hackensack Avenue		
; CITY: Hackensack		
; STATE: New Jersey		
; COUNTRY: USA		
; ZIP: 07601		
; COMPUTER READABLE FORM:		
; MEDIUM TYPE: Floppy disk		
; COMPUTER: IBM PC compatible		
; OPERATING SYSTEM: PC-DOS/MS-DOS		
; SOFTWARE: Patent In Release #1.0, Version #1.25		
; CURRENT APPLICATION DATA:		
; APPLICATION NUMBER: US/08/683,426		
; FILING DATE:		
; CLASSIFICATION: 536		
; PRIOR APPLICATION DATA:		
; APPLICATION NUMBER: 08/312,387		
; FILING DATE: September 26, 1994		
; CLASSIFICATION: 536		
; ATTORNEY/AGENT INFORMATION:		
; NAME: Jackson Esq., David A.		
; REGISTRATION NUMBER: 26,742		
; REFERENCE/DOCKET NUMBER: 600-1-095B		
; TELECOMMUNICATION INFORMATION:		
; TELEPHONE: 201 487-5800		
; TELEFAX: 201 343-1684		
; TELEX: 133521		
; INFORMATION FOR SEQ ID NO: 11:		
; SEQUENCE CHARACTERISTICS:		
; LENGTH: 348 amino acids		
; TYPE: amino acid		
; TOPOLOGY: linear		
; MOLECULE TYPE: protein		
; US-08-683-426-11		
Query Match		4.1%; Score 210; DB 1; Length 348;
Best Local Similarity		26.4%; Pred. No. 4.1e-10;
Matches		77; Conservative 57; Mismatches 110; Indels 48; Gaps 14;
QY	433	PLVSIYPAYNCANYIQCVCDSALNOTVVDLEVCICNDGSDTNTLEVINKLYGNNPRVRI 492
DB	3	PLVSVLICAYNVEKYFAQSLAAVNTWRNLDILIVDDGSTDGTGLAIKDFQRDSRIKI 62
QY	493	MSK-PNGGIASASNAV-SFAK-----GYIQLGSDDDYLEPDAVELCLKEFLKDKTLAC 545
DB	63	LAQAQNSGLIPLSINIGLDELAKSGGGGGEYIARTDADDIASPGWIEKIVGEMEKDRSIIA 122
QY	546	V--YTTNRNVNPDGSLIA-----NGYNWPFERSREKLITAMI-----AHHFRMFTIRAWHLT 594
DB	123	MGAWLEVLSEKDGNRRLARHHKKGKIWKPTRHEDIAAFFFFGNPHNNTMIMRRS--VI 180
QY	595	DG---FNENITENAVDYDMFLKSEVGFKFKHLNKICYNRVHLHGDNTSIK-----KLG 643
DB	181	DGGLRYDTERDAEDYQFWDVSKGLRAYYPEALVKYRLHANQVSSKHSVRQHEIAQGI 240
QY	644	QK--KNHFVVVNSLNROGINNYNDKFDLDESRYKIFNKTAEYQ--EEMDM 692
DB	241	QKTARNDFL---QSMGFK-----TRFDSLE-----YRQTKAAAYELPEKDL 278
RESULT 11		
US-08-683-458-3		
; Sequence 3, Application US/08683458		
; Patent No. 5798233		
; GENERAL INFORMATION:		
; APPLICANT: Gotschlich, Emil C.		
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF		
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM		
; NUMBER OF SEQUENCES: 12		
; CORRESPONDENCE ADDRESS:		
; ADDRESSEE: Klauber & Jackson		
; STREET: 411 Hackensack Avenue		
; CITY: Hackensack		
; STATE: New Jersey		
; COUNTRY: USA		
; ZIP: 07601		
; COMPUTER READABLE FORM:		
; MEDIUM TYPE: Floppy disk		
; COMPUTER: IBM PC compatible		
; OPERATING SYSTEM: PC-DOS/MS-DOS		
; SOFTWARE: Patent In Release #1.0, Version #1.25		
; CURRENT APPLICATION DATA:		
; APPLICATION NUMBER: US/08/683,458		
; FILING DATE:		
; CLASSIFICATION: 435		
; PRIOR APPLICATION DATA:		
; APPLICATION NUMBER: 08/312,387		
; FILING DATE: September 26, 1994		
; CLASSIFICATION: 435		
; ATTORNEY/AGENT INFORMATION:		
; NAME: Jackson Esq., David A.		
; REGISTRATION NUMBER: 26,742		
; REFERENCE/DOCKET NUMBER: 600-1-095A		
; TELECOMMUNICATION INFORMATION:		
; TELEPHONE: 201 487-5800		
; TELEFAX: 201 343-1684		
; TELEX: 133521		
; INFORMATION FOR SEQ ID NO: 3:		
; SEQUENCE CHARACTERISTICS:		
; LENGTH: 348 amino acids		
; TYPE: amino acid		
; TOPOLOGY: linear		
; MOLECULE TYPE: protein		
; US-08-683-458-3		
Query Match		4.1%; Score 210; DB 1; Length 348;
Best Local Similarity		26.4%; Pred. No. 4.1e-10;
Matches		77; Conservative 57; Mismatches 110; Indels 48; Gaps 14;
QY	433	PLVSIYPAYNCANYIQCVCDSALNOTVVDLEVCICNDGSDTNTLEVINKLYGNNPRVRI 492
DB	3	PLVSVLICAYNVEKYFAQSLAAVNTWRNLDILIVDDGSTDGTGLAIKDFQRDSRIKI 62
QY	493	MSK-PNGGIASASNAV-SFAK-----GYIQLGSDDDYLEPDAVELCLKEFLKDKTLAC 545
DB	63	LAQAQNSGLIPLSINIGLDELAKSGGGGGEYIARTDADDIASPGWIEKIVGEMEKDRSIIA 122
QY	546	V--YTTNRNVNPDGSLIA-----NGYNWPFERSREKLITAMI-----AHHFRMFTIRAWHLT 594
DB	123	MGAWLEVLSEKDGNRRLARHHKKGKIWKPTRHEDIAAFFFFGNPHNNTMIMRRS--VI 180
QY	595	DG---FNENITENAVDYDMFLKSEVGFKFKHLNKICYNRVHLHGDNTSIK-----KLG 643
DB	181	DGGLRYDTERDAEDYQFWDVSKGLRAYYPEALVKYRLHANQVSSKHSVRQHEIAQGI 240
QY	644	QK--KNHFVVVNSLNROGINNYNDKFDLDESRYKIFNKTAEYQ--EEMDM 692
DB	241	QKTARNDFL---QSMGFK-----TRFDSLE-----YRQTKAAAYELPEKDL 278

Db 241 QKTARNDL--QSMGFK-----TRFDSLE---YRQTKAAAYELPEKDL 278

Query Match 4.1%; Score 210; DB 1; Length 348;

Query Match	4.1%	Score 210;	DB 2;	Length 348;
Best Local Similarity	26.4%;	Pred. No. 4.1e-110;		
Matches	77;	Conservative	57;	Mismatches 110;
			Indels	48; Gaps 14;
QY	433	PLVSIYIPAYNCANV	IORCVDSALNQTVVDLEVCICDNGSDNTLT	VINKLYGNNPRVRI 492
Db	3	PLVSLICAYNVEYFQASLA	VYVQTVWRNLDIILVDDGSDTGGTATAKD	FOKKRDSRIKI 62
QY	493	MSK-PNGGIASASNAAV-SFAK	-----GYTIGOLDSDDYLEDP	AVELCKLEFKDKKTLLAC 545

Db 63 LAQAQNSGLIPSLINIGLDELAKSGGGGGYIARTDADDIATSPGWIEKIVGEMEKDRSIIA 122

QY 546 V--YTTNRNVNPDGSLIA-----NGYNWPFBSREKLTAMI-----AHHFRMFTIRAWHLT 594

Db 123 MGAWLEVLSEEDGRLARHHKHGKIWKPKTRHEDIAAFPPFCNPIHNTMTIMRRS--VI 180

QY 595 DG---FNENIENAVDVMFLKJSEVGFKHLNKCICNVRVLHGDNSTIK-----KLGI 643

Db 181 DGLRLYDTERDWAEDYQFWYDVSGLGRLAYYPEALVKYRLHANOVSKSHSVROHEIAQGI 240

QY 644 OK--KNHFVVVNSLNROGINNYNDKFDLDESRYIENKTAEO--EEMDM 692

Db 241 OKTARNDFL--QSMGFK-----TRFDSLE-----YRQTKAAAYELPEKDL 278

RESULT 14

US-08-878-360-11 Application US/08878360

; Sequence 11, Patent No. 5945322

; GENERAL INFORMATION:

APPLICANT: Gotschlich, Emil C.

TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF

TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/878,360

FILING DATE: 18-JUN-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/683,426

FILING DATE: 08/312,387

FILING DATE: September 26, 1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-095B

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 348 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-878-360-11

Query Match 4.1%; Score 210; DB 2; Length 348;

Best Local Similarity 26.4%; Pred. No. 4.1e-10;

Matches 77; Conservative 57; Mismatches 110; Indels 48; Gaps 14;

QY 433 PLVSIYIPAYNCANYIQRCDVSDALNQTVDLEVCICNDGSDTNTLVINKLYGNPNRVRI 492

Db 3 PLVSVLCAYNVEKYFAQSLAAVYVNTWRNLILIVDDGSDGTGLAIKDFQKRSRIKI 62

QY 493 MSK-PNGGIASASNAAV-SFAK-----GYVIGQLSDSDYLEPDAVELCLKEFLKDTLAC 545

Db 63 LAQAQNSGLIPSLINIGLDELAKSGGGGGYIARTDADDIATSPGWIEKIVGEMEKDRSIIA 122

QY 546 V--YTTNRNVNPDGSLIA-----NGYNWPFBSREKLTAMI-----AHHFRMFTIRAWHLT 594

Db 123 MGAWLEVLSEEDGRLARHHKHGKIWKPKTRHEDIAAFPPFCNPIHNTMTIMRRS--VI 180

QY 595 DG---FNENIENAVDVMFLKJSEVGFKHLNKCICNVRVLHGDNSTIK-----KLGI 643

Db 181 DGLRLYDTERDWAEDYQFWYDVSGLGRLAYYPEALVKYRLHANOVSKSHSVROHEIAQGI 240

QY 644 OK--KNHFVVVNSLNROGINNYNDKFDLDESRYIENKTAEO--EEMDM 692

Db 241 OKTARNDFL--QSMGFK-----TRFDSLE-----YRQTKAAAYELPEKDL 278

RESULT 15

US-08-478-140B-3

; Sequence 3, Application US/08478140B

; Patent No. 6127153

; GENERAL INFORMATION:

APPLICANT: JOHNSON, KARL F.

APPLICANT: ROTH, STEPHEN

APPLICANT: BUCZALA, STEPHANIE L.

TITLE OF INVENTION: METHOD OF TRANSFERRING AT LEAST TWO

TITLE OF INVENTION: SACCHARIDE UNITS WITH A POLYGLYCOSYLTRANSFERASE, A

TITLE OF INVENTION: POLYGLYCOSYLTRANSFERASE AND GENE ENCODING A

TITLE OF INVENTION: POLYGLYCOSYLTRANSFERASE

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/478,140B

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Laura A. Coruzzi

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7188-017

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-9741/8864

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 348 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-478-140B-3

Query Match 4.1%; Score 210; DB 3; Length 348;

Best Local Similarity 26.4%; Pred. No. 4.1e-10;

Matches 77; Conservative 57; Mismatches 110; Indels 48; Gaps 14;

QY 433 PLVSIYIPAYNCANYIQRCDVSDALNQTVDLEVCICNDGSDTNTLVINKLYGNPNRVRI 492

Db 3 PLVSVLCAYNVEKYFAQSLAAVYVNTWRNLILIVDDGSDGTGLAIKDFQKRSRIKI 62

QY 493 MSK-PNGGIASASNAAV-SFAK-----GYVIGQLSDSDYLEPDAVELCLKEFLKDTLAC 545

Db 63 LAQAQNSGLIPSLINIGLDELAKSGGGGGYIARTDADDIATSPGWIEKIVGEMEKDRSIIA 122

QY 546 V--YTTNRNVNPDGSLIA-----NGYNWPFBSREKLTAMI-----AHHFRMFTIRAWHLT 594

Search completed: January 4, 2003, 02:29:20
Job time : 71 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2003, 02:26:45 ; search time 203 Seconds
(without alignments)
90.090 Million cell updates/sec

Title: US-09-842-484A-2
Perfect score: 5089
Sequence: 1 MNTLSQAIKAYNSNDYELAL.....SAKGENIPVKNFINSITL 965

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 117078 seqs, 18951520 residues
Total number of hits satisfying chosen parameters: 117078

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	486.5	88.2	972	9	US-09-879-959-10
2	276.5	5.4	706	10	Sequence 10, Appl
3	276.5	5.4	715	10	Sequence 4950, Ap
4	246	4.8	322	10	Sequence 10511, A
5	224	4.4	332	10	Sequence 34, Appl
6	223.5	4.4	278	10	Sequence 22, Appl
7	223.5	4.4	674	10	Sequence 36, Appl
8	217.5	4.3	332	10	Sequence 200, App
9	210	4.1	348	12	Sequence 21, Appl
10	210	4.1	348	12	Sequence 3, Appl
11	209	4.1	337	12	Sequence 11, Appl
12	209	4.1	337	12	Sequence 5, Appl
13	208.5	4.1	270	10	Sequence 12, Appl
14	207	4.1	301	10	Sequence 39, Appl
15	205	4.0	389	10	Sequence 27, Appl
16	200.5	3.9	303	10	Sequence 34, Appl
17	199	3.9	297	10	Sequence 29, Appl
18	194	3.8	120	10	Sequence 31, Appl
19	194	3.8	322	10	Sequence 52, Appl
					Sequence 35, Appl

ALIGNMENTS

RESULT 1
US-09-879-959-10
; Sequence 10, Application US/09879959
; Patent No. US20020160489A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H
; APPLICANT: Kumari, Kshama
; APPLICANT: Deangelis, Paul
; TITLE OF INVENTION: STREPTOCOCCUS EQUISIMILIS HYALURONAN SYNTHASE GENE AND EXPRESS
; TITLE OF INVENTION: IN BACILLUS SUBTILIS
; FILE REFERENCE: 3554.049
; CURRENT APPLICATION NUMBER: US/09/879,959
; PRIORITY FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: 09/469,200
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 972
; TYPE: PRT
; ORGANISM: pasteurella multocida
US-09-879-959-10

Query Match	88.2%	Score	4486.5	DB	9	Length	972
Best Local Similarity	86.6%	Pred. No.	0	Mismatches	60	Indels	7
Matches	842	Conservative	63	Gaps	2		
Qy	1	MNTLSQAIKAYNSNDYELALKLFKSAEYGRKIVFQIKCKEL	---	STNS	---	YVS	53
Db	1	MNTLSQAIKAYNSNDYELALKLFKSAEYGRKIVFQIKCKEL	---	STNS	---	YVS	53
Qy	54	EDKKNVCSDDIATQLLSNVYKLTLSSEKNSLKNKWSITGKKSNAEIRKVELP	113				
Db	61	KEEVNVCSDPLDIATQLLSNVYKLTLSSEKNSLKNKWSITGKKSNAEIRKVELP	120				
Qy	114	KDFPKDLVLAFLPDHVNDFTWYKRNKKSGLGKPVKNIGLSIIPTFNSRILDIACL	173				
Db	121	KDFPKDLVLAFLPDHVNDFTWYKRNKKSGLGKPVKNIGLSIIPTFNSRILDIACL	180				
Qy	174	VNQKNTYPEVVVADGSKENLLTIVQKVEQKLDIKYVRQKDYQYLCVAVRNGLRTAKY	233				

QY 708 VSIFYPNTLNGLVKKLNIIYV--KNIFVIIHVDKNHLT--PDIK-KEILAFYKHQV 762
DB 255 -----VKCLNKMYSTDCLDNEFLPILESYRKEIRYFFIKARYLS--RKHLV 300
QY 763 NILL 766
DB 301 TLVL 304
RESULT 5
US-09-767-041-22
; Sequence 22, Application US/09767041
; Patent No. US20020055168A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Hilda
; TITLE OF INVENTION: STREPTOCOCCUS SUI S VACCINES AND DIAGNOSTIC TESTS
; FILE REFERENCE: 2183-4726
; CURRENT APPLICATION NUMBER: US/09/767,041
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: PCT/NL99/00460
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: EP98202465.5
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: EP98202467.1
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 22
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Streptococcus suis
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: CPS2K
US-09-767-041-22

Query Match 4.4%; Score 224; DB 10; Length 332;
Best Local Similarity 24.4%; Pred. No. 3.1e-09;
Matches 81; Conservative 65; Mismatches 118; Indels 68; Gaps 14;
QY 435 VSIYIPAYNCANYIQRVDSALNQTVDVLEVCICNDGSDTNTLEVINKLKYNPNRVRIMS 494
DB 4 ISIIIVPNVBOYLKSCINSIVNQTYSKHIEILLVNDGSDTNSDSEICLAYAKKDSRIRYFK 63
QY 495 KPNGGSIASNAVSFAKGYIGOLDSDYLPDAVELCLKEFLKDKTKLACVYITNR--- 551
DB 64 KENGGLSDARNYGISRAKGDIYAFIDSDDFIHSEFIQRLHEAIBERENALVAVAGYDRVDA 123
QY 552 -----NVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHLTDGFEINENIA 604
DB 124 SGHFLTAELPTNQAVLSGRN---VCKKLEA-DGHRF-----VVAWN-----KL 164
QY 605 VDYDMFLKSEVGKPKHLNKYCNKVLHGDNTSKKGIQKKNHFVVVQNQSLNRGINYY 664
DB 165 YKKELDFRFEKGKI-HEDEYFTYRLLY-----ELEKVAIVKE-----CLYI 205
QY 665 NYDKFDLDES--RKYIFNKTAQYQ-BEMDML-----KDLKLIQNKDAKIAVSIF---Y 712
DB 206 YVDRENSIITSMIDHREHFCLEFQNERMDYFESRGDEL-LLECYSRFLAFVFLGKY 264
QY 713 PNTLNGLVKLNLI-----IEYNKNIFVII 737
DB 265 NHWLSKQOKKLQTLFRIVYKQLKQNKRLALLM 296

RESULT 6
US-09-767-041-36
; Sequence 36, Application US/09767041
; Patent No. US20020055168A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Hilda
; TITLE OF INVENTION: STREPTOCOCCUS SUI S VACCINES AND DIAGNOSTIC TESTS

FILE REFERENCE: 2183-4726
CURRENT APPLICATION NUMBER: US/09/767,041
CURRENT FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: PCT/NL99/00460
PRIOR FILING DATE: 1999-07-19
PRIOR APPLICATION NUMBER: EP98202465.5
PRIOR FILING DATE: 1998-07-22
PRIOR APPLICATION NUMBER: EP98202467.1
PRIOR FILING DATE: 1998-07-22
NUMBER OF SEQ ID NOS: 53
SOFTWARE: Patentin version 3.0
SEQ ID NO 36
LENGTH: 278
TYPE: PRT
ORGANISM: Streptococcus suis
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: CPS1K
US-09-767-041-36
Query Match 4.4%; Score 223.5; DB 10; Length 278;
Best Local Similarity 34.1%; Pred. No. 2.7e-09;
Matches 44; Conservative 29; Mismatches 55; Indels 1; Gaps 1;
QY 432 IPLVSIYIPAYNCANYIQRVDSALNQTVDVLEVCICNDGSDTNTLEVINKLKYNPNRVR 491
DB 4 ISKISIIIVPNVBEKYLKSCIDSIYNQYKHIEILLVNDGSDTNSDSEICLAYAKKDSRIR 63
QY 492 IMSKPNGGSIASNAVSFAKGYIGOLDSDYLPDAVELCLKEFLKDKTKLACVYITNR 551
DB 64 YFKKENGGLSDARNYGISRAKGDIYAFIDSDDFIHSEFIQRLHEAIBERENALVAVAGYDR 123
QY 552 NVNPDGSLI 560
DB 124 -VDASGHFL 131
RESULT 7
US-09-765-272-200
; Sequence 200, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Chol et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 200:
; SEQUENCE CHARACTERISTICS:

LENGTH: 674 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 200:
US-09-765-272-200

Query Match 4.4%; Score 223.5; DB 10; Length 674;
Best Local Similarity 20.1%; Pred. No. 9.1e-09;
Matches 96; Conservative 89; Mismatches 177; Indels 115; Gaps 17;

Qy 451 CVDASALNQTVDVLEVCICNDSTNTLEVINLKYNNPRVRIMSPNGGIASASNAAYSF 510
Db 1 CLDSIITQYKNIETIVVNDGSDASGEICKEFSEMDHRILYIEQENAGLSAARTGLNN 60
Qy 511 AKGYIGOLDSDYLEPAVELCLK---EFLKDKTLACVYTTNRN-----552
Db 61 MSGNVTFTVDSDDTEQDYVETLYKKIVEYQADIAVGNYYSFNESEGMFYPHILGDSYFE 120
Qy 553 -VNPDGSLIANGYNWPEFSREKLTAMIAHHRMFTIRAWHL-----TDGFENIE 602
Db 121 KYDWNVSIFENLYE---TQEMKSFALISAWGKLYKARLFEQLRFDIGKLGEDGY-----171
Qy 603 NAVDYDMFLKSEVGKFKHLNKICY-NRVLHGD-----NTSIRKL 641
Db 172 --LNQKVYL-LSE--KVIYLNKSLYAIRKGSLSRVWTEKWMHALVDMASERITLLANM 226
Qy 642 GIQKNHFVVNQSLNRQGINYYNDKFDLDESRYIFENKTAEOEMDKLKLQN 701
Db 227 GYPLEKHLAVRQMLEVSLAN---GOASGLSDTATY-----KEPEMKORLLNQLSROBE 277
Qy 702 KDAIAVSIFVPNTLNGVLKLNIEYNKIFVILHVD-KNHLTPDIKKEILAF----756
Db 278 SEKKAIVLAANYGYVDVLTIKSCYHNRSIRFLIHSDFNEWIKQLNKKLEFDSFI 337
Qy 757 ----YHKHQVNILNNDISYTSNRKIRKTEAHLNKLQNLNCEYIIFDNHDSLFVK 812
Db 338 INCRVTSEQIS-CYKSDISYTVFLRYFTADF----VQEDKALYLCDLVVTNKLDDLFAT 392
Qy 813 N-DSY-----AYWKYDVGCMNFSALTHDWEIKIN 840
Db 393 DLQDYPLAARDFGGRAYGQEIFNAGVLLVNNAPMKKENMTQKLDIVTNEWHDKVD 449

RESULT 8
US-09-767-041-21
; Sequence 21, Application US/09767041
; Patent No. US20020055168A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Hilda
; TITLE OF INVENTION: STREPTOCOCCUS SUITS VACCINES AND DIAGNOSTIC TESTS
; FILE REFERENCE: 2183-4726
; CURRENT APPLICATION NUMBER: US/09/767,041
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: PCT/NL99/00460
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: EP98202465.5
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: EP98202467.1
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 21
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Streptococcus suis
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: CFS2J
US-09-767-041-21

Query Match 4.3%; Score 217.5; DB 10; Length 332;

Best Local Similarity 24.2%; Pred. No. 9.4e-09;
Matches 86; Conservative 67; Mismatches 116; Indels 87; Gaps 18;
Qy 435 VSIYIPAYNCANYIQRVCDSALNQTVDVLEVCICNDSTNTLEVINLKYNNPRVRIMS 494
Db 4 VSIIVPIFNTKYLRECLDSIIISQSYTNLEILLIDGSDSDTDICLEYAEOQDRIKFLR 63
Qy 495 .KPNGGIASASNAAVSFAGKYYIGOLDSDYLEPDAVE---LGLKEFLKDKTLACVYTTNR 551
Db 64 LPNGGVSARNYGIKNTANYIMFVDSDDIVDGNIVESLYTCLKE--NSDLS-----114
Qy 552 NVNPDGSLIAN-GYNWPEFSREK-----LTTAMIAHHRF--MFT-----IR 589
Db 115 -----GGLLATFDGNYQESELQCKIDLEIKERDGLNENPNHYMSGIFNSPCKLYK 169
Qy 590 AWHLDGTFNENENAVDYDMFLKLS---EVGKFKHLNKICY--NRVLHGDNTSIK-KLGI 643
Db 170 NYIINGGF--DTEQWLGEDLLENLYKNKIKVRYVNRNLXFARRSLQSTTNTFKYDVFI 227
Qy 644 QKKN-----HFVVVNQSLNRQGINYYNDKFDLDES-----KYI 679
Db 228 QLENLEEKTFDLFVKIFGQYEFVFKETL-QWHIIYYSLLMEKNGESLPKKLHIFKYL 286
Qy 680 FNKTAEOEMDML--KDLKLIONDAKTAVSIFYPTNTLNLGVK-KLNIIIEYNKN 732
Db 287 YNR-----HSLDTLSIKRTSSVFKRICKLIVA-----NNLFKIFLNTLIREKN 330

RESULT 9
US-10-007-267-3
; Sequence 3, Application US/10007267
; Patent No. US20020127682A1
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0. Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/007,267
; FILING DATE: 03-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/333,412
; FILING DATE: 15-Jun-1999
; APPLICATION NUMBER: 08/312,387
; FILING DATE: July 7, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-10-007-267-11

Query Match 4.1%; Score 210; DB 12; Length 348;
Best Local Similarity 26.4%; Pred. No. 3.6e-08;
Matches 77; Conservative 57; Mismatches 110; Indels 48; Gaps 14;

QY 433 PLSYIIPAYNCANYTORCVDSALNOTVVDLEVCINCNDGSDTNTLEVINKLYGNPRVRI 492
DB 3 PLSVLICAYNVEKYFAQSLAAVNOTWRNLDLIYVDDGSDTGTDLAIKDFQRDSRIKI 62
QY 493 MSK-PNGGIASNAAV-SFAK-----GYIGOLDSDDYLEPDAVELCLKEFLKDKTLAC 545
DB 63 LAQAQNSGLIPLSNIGLDELAKSGGGGEYIARTDADDIASPGWIEKIVGEMEKDRSIIA 122
QY 546 V--YTTNRNVNPDGSLIA-----NGYNWPEFSREKLTAMI-----AHHFRMFTIRAWHLT 594
DB 123 MGAWLEVLSEKDGRLARHHKHKWKPTRHEDIAAFAFPFGNPHNNTMIMRRS--VI 180
QY 595 DG---FNENTENAVDVMFLKSEVGFKHLNKICYNRVLHGDNTSIK-----KLGI 643
DB 181 DGLRYDTERDWAEDYQFWYDVSKLGLAYYPEALVKYRLHANOVSSKHSVRQHEIAQGI 240
QY 644 QK--KNHFVVVNSLNROGINNYNDKFDLDESRYKFIKNTAEYQ--EEMDM 692
DB 241 QKTARNDL---QSMGFK-----TRFDSLE-----YRQTKAAAYELPEKDL 278

RESULT 10

US-10-007-267-11
Sequence 11, Application US/10007267
Patent No. US20020127682A1

GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND GENES ENCODING THEM

NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/007,267
FILING DATE: 03-Dec-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/333,412
FILING DATE: 15-Jun-1999
APPLICATION NUMBER: 08/312,387
FILING DATE: July 7, 1994

ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095

TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521

INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-007-267-11

US-10-007-267-5

Query Match 4.1%; Score 210; DB 12; Length 348;
Best Local Similarity 26.4%; Pred. No. 3.6e-08;
Matches 77; Conservative 57; Mismatches 110; Indels 48; Gaps 14;

QY 433 PLSYIIPAYNCANYTORCVDSALNOTVVDLEVCINCNDGSDTNTLEVINKLYGNPRVRI 492
DB 3 PLSVLICAYNVEKYFAQSLAAVNOTWRNLDLIYVDDGSDTGTDLAIKDFQRDSRIKI 62
QY 493 MSK-PNGGIASNAAV-SFAK-----GYIGOLDSDDYLEPDAVELCLKEFLKDKTLAC 545
DB 63 LAQAQNSGLIPLSNIGLDELAKSGGGGEYIARTDADDIASPGWIEKIVGEMEKDRSIIA 122
QY 546 V--YTTNRNVNPDGSLIA-----NGYNWPEFSREKLTAMI-----AHHFRMFTIRAWHLT 594
DB 123 MGAWLEVLSEKDGRLARHHKHKWKPTRHEDIAAFAFPFGNPHNNTMIMRRS--VI 180
QY 595 DG---FNENTENAVDVMFLKSEVGFKHLNKICYNRVLHGDNTSIK-----KLGI 643
DB 181 DGLRYDTERDWAEDYQFWYDVSKLGLAYYPEALVKYRLHANOVSSKHSVRQHEIAQGI 240
QY 644 QK--KNHFVVVNSLNROGINNYNDKFDLDESRYKFIKNTAEYQ--EEMDM 692
DB 241 QKTARNDL---QSMGFK-----TRFDSLE-----YRQTKAAAYELPEKDL 278

RESULT 11

US-10-007-267-5
Sequence 5, Application US/10007267
Patent No. US20020127682A1

GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND GENES ENCODING THEM

NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/007,267
FILING DATE: 03-Dec-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/333,412
FILING DATE: 15-Jun-1999
APPLICATION NUMBER: 08/312,387
FILING DATE: July 7, 1994

ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095

TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 337 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-007-267-5

Db 117 -----DAFVHRVTKQYRFKQDEVFNQKELEFLSKQRH-----FCWSVWAKCFKKDI 165
QY 610 FLKLSVGKF-KHLN-----KICYNRVLHGDNTSIKKLGIOKKNHFFVVVQNROGINY 663
Db 166 ILKSEFKIKIDERLNYGEDVLFYIYFMCEKIAVFTKCI----- 205
QY 664 YNDKFDLDESRYIFNKTAEQEEMDKLKLQNDKAKTAVSIFY-PNTLNGLVKK 722
Db 206 YHYE-----FNPNGRY-----ENKKEILNQNYHDKKKSNEIILK 240
QY 723 LNNII---EYNKNIFVILLHVD---KNHL 745
Db 241 LSKEFAHDEPHOKLFEVLKKEEAGVKNRL 269
RESULT 14
US-09-816-028A-27
; Sequence 27, Application US/09816028A
; Patent No. US20020042369A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816, 028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 27
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: beta-1,3-galactosyltransferase from C. jejuni strain
; OTHER INFORMATION: OH4384 (ORF 6a of lipooligosaccharide (LOS)
; OTHER INFORMATION: biosynthesis locus)
us-09-816-028A-27
Query Match 4.18; Score 207; DB 10; Length 301;
Best Local Similarity 21.08; Pred. No. 4.9e-08;
Matches 78; Conservative 64; Mismatches 116; Indels 114; Gaps 13;
QY 435 VSIYIPAYNCANYIQRCDVSALNTQVVDLEVCICNDGSTDNTLEVINKLYGNPRVRIM- 493
Db 4 ISIILPTVNVQYIARATESCINQTFKDIETIIVDDCGNDNSINIAKEYSKDKRIKIIH 63
QY 494 SKPNGGIASANAASFAGKYIIGQLSDDDYLEPDAVELCKEFLKDKTLACVYTTNRV 553
Db 64 NEKNLGLRLARYEGYKVANSPTIIMFLDPDDYLELNACCEKILDEQDEVDLVF----- 117
QY 554 NPDGSLIANGYNWPEFSREKLTMTAMIAHFRMTIRAWHLTDGPNTE-NAVYDYMELK 612
Db 118 -----KKFDNPSGFSYKKEPVKKI-TAKKNLY-----FNAIVESNVISY----- 129
QY 613 LSEVGKFKHLNKICYNRVLHGDNTSIKKLGIOKKNHFFVVVQNROGINYNDKFDL 672
Db 130 -----KKFDNPSGFSYKKEPVKKI-TAKKNLY-----WTMWSGL 162
QY 673 DESRYIFNKTAEQEEMDKLKLQNDKAKIAVSIFYPNTLNGLVKLNIIYEYKN 732
Db 163 IRKKLYL-----EAPASFLERKOVKINMAED-----VLIYYP-----MLSQAKOIAMN 208
QY 733 IFVILLHVDKHNLTDPDIKKEILAFYHKHQVNILLNNDISYTSNRLIKTEAHLNINK-L 791
Db 209 LY---HYVPNNNSICNTKNEVLVKNIIQELQLVLN---YLRQNYI-----LNKYC 252

QY 792 SOLNLNCEYIIF 803
Db 253 SVLYVLIKVLLY 264
RESULT 15
US-09-816-028A-34
; Sequence 34, Application US/09816028A
; Patent No. US20020042369A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816, 028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 34
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: glycosyltransferase of C. jejuni OH4384 (ORF 4a of
; OTHER INFORMATION: lipooligosaccharide (LOS) biosynthesis locus)
us-09-816-028A-34
Query Match 4.08; Score 205; DB 10; Length 389;
Best Local Similarity 21.28; Pred. No. 9.9e-08;
Matches 114; Conservative 81; Mismatches 146; Indels 196; Gaps 25;
QY 435 VSIYIPAYNCANYIQRCDVSALNTQVVDLEVCICNDGSTDNTLEVINKLYGNPRVRIM 493
Db 4 IGVYIPIYNVKYLRECLDSVINQTYNLEIILVNDGSTDDEHSLNIAKEYTLKDKRIITLF 63
QY 494 SKPNGGIASANAASFAGKYIIGQLSDDDYLEPDAVELCKEFLKDKTLACVYTTNRV 553
Db 64 DKNKGGLSSARNIGIEYFSGEY-----KLKNTQHIKENSLEIF 102
QY 554 NPDGSLIANGYNWPEFSREKLTMTAMIAHFRMTIRAWHLTDGPN-ENIEN----AVDY 607
Db 103 QLDGN---NPYN-----IYKAYKSSQAFNNEKDLTNFTYPSIDY 138
QY 608 DMFLKLSVGKFKHLNKICYNRVLHGDNTSIKKLGIOKKNHFFVVVQNROGINYND 667
Db 139 IIFLSDSNYWL-----NCIEECVIRKKNVDVL-----WFDHD 171
QY 668 -KFDLDESRYIFNKTAEQEEM-DMLKDLKLQNDKAKIAVSIFYPNTLNGLVKLN 725
Db 172 CTYED-----NKNKKHKTKMEIFDFKECIITPEYANRALSVGRDISFG----WNG 221
QY 726 IIEYNKNIFVILLHVDKHNLTDPDIKKEILAFYHKHQVNILLNNDIS-----YTSNRLIK 780
Db 222 MIDFN-----FLQIKLKF-----INFIINEDIHFGILFASANKIYV 259
QY 781 TEAHLNINKLSQNLNCEYIIFDNDHDSLFVKNDYVAYMKYDVGVMNFSALTDHTEKIN 840
Db 260 ----LSQKLYLCRLRANS-----ISNHD-----ITKAN 285
QY 841 AHPPPKLIKITYFNDNDRSMN-VKASQSGMPMKYALPHELITIIKEVITSCOSIDSYPE 899
Db 286 VSEYFKDIYET-FGENAKEAKNLYKAASR-----VITALKLIIEFFKD 336
QY 900 YNTEDWFOFALLIIEKTKGHVFNFTSTUTYMPWPKLQWTFNEQIOSAKKGENIPVN 956
Db 327 QKNEN-----ALAIKE-----TFLPCYAK-----KALMIKKFKKDKPLN 359

Search completed: January 4, 2003, 02:49:25
Job time : 206 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2003, 02:23:10 ; Search time 400 Seconds
(without alignments)
1555.419 Million cell updates/sec

Title: US-09-842-484A-2

Perfect score: 5089

Sequence: 1 MNTLSQAIKAYNSNDYELAL.....SAKKGENIPVKNFIINSITL 965

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

```
1: /cgn2_6/ptodata/1/paa/US06_COMB.pcp.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pcp.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pcp.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pcp.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pcp.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pcp.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pcp.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pcp.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pcp.*
10: /cgn2_6/ptodata/1/paa/US085_COMB.pcp.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pcp.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pcp.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pcp.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pcp.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pcp.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pcp.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pcp.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pcp.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pcp.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pcp.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pcp.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pcp.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pcp.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pcp.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pcp.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pcp.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pcp.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5089	100.0	965	1 PCT-US01-13395-2	Sequence 2, Appli
2	5089	100.0	965	1 PCT-US99-26501-3	Sequence 3, Appli
3	5067	99.6	965	1 PCT-US01-13395-4	Sequence 4, Appli
4	5017.5	98.6	972	1 PCT-US99-07289-3	Sequence 3, Appli
5	5017.5	98.6	972	1 PCT-US99-07289-8	Sequence 8, Appli
6	5017.5	98.6	972	16 US-09-283-402-3	Sequence 3, Appli

7	5017.5	98.6	972	16	US-09-283-402-8	Sequence 8, Appli
8	5017.5	98.6	972	18	US-09-469-200-3	Sequence 3, Appli
9	5017.5	98.6	972	18	US-09-469-200-8	Sequence 8, Appli
10	4601.5	90.4	972	1	PCT-US99-07289-9	Sequence 9, Appli
11	4601.5	90.4	972	16	US-09-283-402-9	Sequence 9, Appli
12	4601.5	90.4	972	18	US-09-469-200-9	Sequence 9, Appli
13	4493.5	88.3	972	21	US-09-791-537-56855	Sequence 56855, A
14	4490.5	88.2	972	24	US-10-011-768B-9	Sequence 9, Appli
15	4490.5	88.2	972	24	US-10-011-771B-9	Sequence 9, Appli
16	4486.5	88.2	972	1	PCT-US99-07289-1	Sequence 1, Appli
17	4486.5	88.2	972	16	US-09-283-402-1	Sequence 1, Appli
18	4486.5	88.2	972	18	US-09-469-200-1	Sequence 1, Appli
19	4486.5	88.2	972	21	US-09-791-537-118183	Sequence 118183, A
20	4486.5	88.2	972	22	US-09-879-959-10	Sequence 10, Appli
21	4486.5	88.2	972	25	US-10-172-527-10	Sequence 10, Appli
22	4479.5	88.0	972	1	PCT-US99-07289-7	Sequence 7, Appli
23	4479.5	88.0	972	16	US-09-283-402-7	Sequence 7, Appli
24	4479.5	88.0	972	18	US-09-469-200-7	Sequence 7, Appli
25	4476.5	88.0	972	1	PCT-US01-13395-8	Sequence 8, Appli
26	3079	60.5	702	1	PCT-US99-26501-1	Sequence 1, Appli
27	523	10.3	101	1	PCT-US99-07289-10	Sequence 10, Appli
28	523	10.3	101	16	US-09-283-402-10	Sequence 10, Appli
29	523	10.3	101	18	US-09-469-200-10	Sequence 10, Appli
30	333.5	6.6	842	15	US-09-134-000-5749	Sequence 5749, Ap
31	299	5.9	331	20	US-09-634-238-247	Sequence 247, App
32	292	5.7	721	15	US-09-107-532-6889	Sequence 6889, Ap
33	292	5.7	721	15	US-09-107-532A-6889	Sequence 6889, Ap
34	277.5	5.5	285	22	US-09-897-516-5714	Sequence 5714, Ap
35	277.5	5.5	285	27	US-60-215-161-5714	Sequence 5714, Ap
36	276.5	5.4	706	1	PCT-US02-03987-4950	Sequence 4950, Ap
37	276.5	5.4	706	21	US-09-791-537-117740	Sequence 117740, A
38	276.5	5.4	706	22	US-09-815-242-4950	Sequence 4950, Ap
39	276.5	5.4	706	24	US-10-072-851-4950	Sequence 4950, Ap
40	276.5	5.4	706	27	US-60-269-308-4735	Sequence 4735, Ap
41	276.5	5.4	715	1	PCT-US02-03987-10511	Sequence 10511, A
42	276.5	5.4	715	22	US-09-815-242-10511	Sequence 10511, A
43	276.5	5.4	715	24	US-10-072-851-10511	Sequence 10511, A
44	262.5	5.2	534	15	US-09-134-000-5087	Sequence 5087, Ap
45	259.5	5.1	323	21	US-09-791-537-16847	Sequence 16847, A

ALIGNMENTS

RESULT 1

PCT-US01-13395-2

; Sequence 2, Application PC/TUS0113395

; GENERAL INFORMATION:

; APPLICANT: DE ANGELIS, PAUL L.

; TITLE OF INVENTION: CHONDROITIN SYNTHASE GENE AND METHODS OF MAKING AND

; TITLE OF INVENTION: USING SAME

; FILE REFERENCE: 618755-9/JP/199,538

; CURRENT APPLICATION NUMBER: PCT/US01/13395

; CURRENT FILING DATE: 2001-04-25

; PRIOR APPLICATION NUMBER: 09/437,277

; PRIOR FILING DATE: 1999-11-10

; PRIOR APPLICATION NUMBER: 60/199,538

; PRIOR FILING DATE: 2000-04-25

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 965

; TYPE: PRT

; ORGANISM: Pasteurella multocida

PCT-US01-13395-2

Query Match 100.0%; Score 5089; DB 1; Length 965;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 965; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGKIVFQIIKCKELSTNSYVSEDKKNSV 60

|||||

Db 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGKIVFQIIKCKELSTNSYVSEDKKNSV 60

|||||

```

; CURRENT APPLICATION NUMBER: PCT/US99/26501
; EARLIER FILING DATE: 1999-11-10
; EARLIER APPLICATION NUMBER: 60/107,929
; EARLIER FILING DATE: 1998-11-11
; EARLIER APPLICATION NUMBER: 09/283,402
; EARLIER FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 965
; TYPE: PRT
; ORGANISM: Pasteurella multocida
; PCT-US99-26501-3

Query Match      100.0%; Score 5089; DB 1; Length 965;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 965; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 CDSLDIATQLLSNVKKTITLSESEKNSLKNKWSITGKKSSENAEIRKVELVPKDFPKDL 120
Db 61 CDSLDIATQLLSNVKKTITLSESEKNSLKNKWSITGKKSSENAEIRKVELVPKDFPKDL 120
QY 121 VLAPLDPHVNDFTWYKRNKKSIGIKPVNKNIGLSIIIPFNRSRIIDITLACLVNOKTNY 180
Db 121 VLAPLDPHVNDFTWYKRNKKSIGIKPVNKNIGLSIIIPFNRSRIIDITLACLVNOKTNY 180
QY 181 PFEVVVADGSGKENLLTIVQKYEQLDIKYVRQDYGQYQLCAVRNLGLRTAKYDFVSILD 240
Db 181 PFEVVVADGSGKENLLTIVQKYEQLDIKYVRQDYGQYQLCAVRNLGLRTAKYDFVSILD 240
QY 241 CDMAPOOLWVHSYLTLELLEDNDIVLIGPRKYVDTHNITAEQFLNDPYYLIESLPETATNN 300
Db 241 CDMAPOOLWVHSYLTLELLEDNDIVLIGPRKYVDTHNITAEQFLNDPYYLIESLPETATNN 300
QY 301 PSITSKGNISLDRLEHFKKTDNLRCDSPFRYFVAGNVAFSKEWLNKVGWDFEENHWG 360
Db 301 PSITSKGNISLDRLEHFKKTDNLRCDSPFRYFVAGNVAFSKEWLNKVGWDFEENHWG 360
QY 361 GEDVEFGYRLFAKGCFFRVIDGGMAIHOEPPGKENETEREAGKSITLKIYKEKVPYIYRK 420
Db 361 GEDVEFGYRLFAKGCFFRVIDGGMAIHOEPPGKENETEREAGKSITLKIYKEKVPYIYRK 420
QY 421 LLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVDVLEVCICNDGSTDNTLEVI 480
Db 421 LLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVDVLEVCICNDGSTDNTLEVI 480
QY 481 NKLYGNPRVRIMSKPNGGSIASNAAVSFAKGYIIGQLSDDDYLEPDAVELCLKEFLKD 540
Db 481 NKLYGNPRVRIMSKPNGGSIASNAAVSFAKGYIIGQLSDDDYLEPDAVELCLKEFLKD 540
QY 541 KTLACVYTTNRNVPDGSILIANGYNWPEFSREKLTAMIAHHFRMFTIRAWHLTDGFEN 600
Db 541 KTLACVYTTNRNVPDGSILIANGYNWPEFSREKLTAMIAHHFRMFTIRAWHLTDGFEN 600
QY 601 IENAVDYDMFLKLESEVGKFKHLNLCYNRVLHGDNTSIKKGIOKKNHVVVNGSLNRQ 660
Db 601 IENAVDYDMFLKLESEVGKFKHLNLCYNRVLHGDNTSIKKGIOKKNHVVVNGSLNRQ 660
QY 661 INYNYDKFDLDESRRKIFNKTAEQEEMDMKDLKLIQNKDIAKTAVSIFYPNTLNGLV 720
Db 661 INYNYDKFDLDESRRKIFNKTAEQEEMDMKDLKLIQNKDIAKTAVSIFYPNTLNGLV 720
QY 721 KKLNNIIEYNKNIFVILHVDKNHLPDIIKKEILAFYHKHQNILLNNDISYTSNRLIK 780
Db 721 KKLNNIIEYNKNIFVILHVDKNHLPDIIKKEILAFYHKHQNILLNNDISYTSNRLIK 780
QY 781 TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMNFSALTHDWEIKIN 840
Db 781 TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMNFSALTHDWEIKIN 840
```

```

; CURRENT APPLICATION NUMBER: PCT/US99/26501
; EARLIER FILING DATE: 1999-11-10
; EARLIER APPLICATION NUMBER: 60/107,929
; EARLIER FILING DATE: 1998-11-11
; EARLIER APPLICATION NUMBER: 09/283,402
; EARLIER FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 965
; TYPE: PRT
; ORGANISM: Pasteurella multocida
; PCT-US99-26501-3

Query Match      100.0%; Score 5089; DB 1; Length 965;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 965; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNTLSQAIKAYNSNDYELALKLFPEKSAETYGKRIVEFQIIKCKEKLSTNSVSEDKKNSV 60
Db 1 MNTLSQAIKAYNSNDYELALKLFPEKSAETYGKRIVEFQIIKCKEKLSTNSVSEDKKNSV 60
QY 61 CDSLDIATQLLSNVKKTITLSESEKNSLKNKWSITGKKSSENAEIRKVELVPKDFPKDL 120
Db 61 CDSLDIATQLLSNVKKTITLSESEKNSLKNKWSITGKKSSENAEIRKVELVPKDFPKDL 120
QY 121 VLAPLDPHVNDFTWYKRNKKSIGIKPVNKNIGLSIIIPFNRSRIIDITLACLVNOKTNY 180
Db 121 VLAPLDPHVNDFTWYKRNKKSIGIKPVNKNIGLSIIIPFNRSRIIDITLACLVNOKTNY 180
QY 181 PFEVVVADGSGKENLLTIVQKYEQLDIKYVRQDYGQYQLCAVRNLGLRTAKYDFVSILD 240
Db 181 PFEVVVADGSGKENLLTIVQKYEQLDIKYVRQDYGQYQLCAVRNLGLRTAKYDFVSILD 240
QY 241 CDMAPOOLWVHSYLTLELLEDNDIVLIGPRKYVDTHNITAEQFLNDPYYLIESLPETATNN 300
Db 241 CDMAPOOLWVHSYLTLELLEDNDIVLIGPRKYVDTHNITAEQFLNDPYYLIESLPETATNN 300
QY 301 PSITSKGNISLDRLEHFKKTDNLRCDSPFRYFVAGNVAFSKEWLNKVGWDFEENHWG 360
Db 301 PSITSKGNISLDRLEHFKKTDNLRCDSPFRYFVAGNVAFSKEWLNKVGWDFEENHWG 360
QY 361 GEDVEFGYRLFAKGCFFRVIDGGMAIHOEPPGKENETEREAGKSITLKIYKEKVPYIYRK 420
Db 361 GEDVEFGYRLFAKGCFFRVIDGGMAIHOEPPGKENETEREAGKSITLKIYKEKVPYIYRK 420
QY 421 LLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVDVLEVCICNDGSTDNTLEVI 480
Db 421 LLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVDVLEVCICNDGSTDNTLEVI 480
QY 481 NKLYGNPRVRIMSKPNGGSIASNAAVSFAKGYIIGQLSDDDYLEPDAVELCLKEFLKD 540
Db 481 NKLYGNPRVRIMSKPNGGSIASNAAVSFAKGYIIGQLSDDDYLEPDAVELCLKEFLKD 540
QY 541 KTLACVYTTNRNVPDGSILIANGYNWPEFSREKLTAMIAHHFRMFTIRAWHLTDGFEN 600
Db 541 KTLACVYTTNRNVPDGSILIANGYNWPEFSREKLTAMIAHHFRMFTIRAWHLTDGFEN 600
QY 601 IENAVDYDMFLKLESEVGKFKHLNLCYNRVLHGDNTSIKKGIOKKNHVVVNGSLNRQ 660
Db 601 IENAVDYDMFLKLESEVGKFKHLNLCYNRVLHGDNTSIKKGIOKKNHVVVNGSLNRQ 660
QY 661 INYNYDKFDLDESRRKIFNKTAEQEEMDMKDLKLIQNKDIAKTAVSIFYPNTLNGLV 720
Db 661 INYNYDKFDLDESRRKIFNKTAEQEEMDMKDLKLIQNKDIAKTAVSIFYPNTLNGLV 720
QY 721 KKLNNIIEYNKNIFVILHVDKNHLPDIIKKEILAFYHKHQNILLNNDISYTSNRLIK 780
Db 721 KKLNNIIEYNKNIFVILHVDKNHLPDIIKKEILAFYHKHQNILLNNDISYTSNRLIK 780
QY 781 TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMNFSALTHDWEIKIN 840
Db 781 TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMNFSALTHDWEIKIN 840
```

RESULT 2
PCT-US99-26501-3
; Sequence 3, Application PC/TUS9926501
; GENERAL INFORMATION:
; APPLICANT: THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA
; TITLE OF INVENTION: POLYMER GRAFTING BY POLYSACCHARIDE SYNTHASES
; FILE REFERENCE: 617481-5

Qy 841 AHPPEKLIKTYFNDNLSNMVKGASQGMFKYALPHELLTIIKEVITSCQIDSVPY 900
Db 841 AHPPEKLIKTYFNDNLSNMVKGASQGMFKYALPHELLTIIKEVITSCQIDSVPY 900
Qy 901 NTEDIWFQFALLILEKKTGHVFNKSTSTLTYPWPKLQWNTNEQIOSAKKGENIPVNFII 960
Db 901 NTEDIWFQFALLILEKKTGHVFNKSTSTLTYPWPKLQWNTNEQIOSAKKGENIPVNFII 960
Qy 961 NSITL 965
Db 961 NSITL 965
RESULT 3
PCT-US01-13395-4
; Sequence 4, Application PC/TUS0113395
; GENERAL INFORMATION:
; APPLICANT: DE ANGELIS, PAUL L.
; TITLE OF INVENTION: CHONDROITIN SYNTHASE GENE AND METHODS OF MAKING AND
; FILE REFERENCE: 618755-9/JP/199,538
; CURRENT APPLICATION NUMBER: PCT/US01/13395
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 09/437,277
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 965
; TYPE: PRT
; ORGANISM: Pasteurella multocida
PCT-US01-13395-4

Query Match 99.6%; Score 5067; DB 1; Length 965;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 960; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy 1 MNTLSQAIKAYNSNDYELALKLFKSAETGRKIVFQIIKCKEKLSTNSYVSEDKKNSV 60
Db 1 MNTLSQAIKAYNSNDYELALKLFKSAETGRKIVFQIIKCKEKLSTNSYVSEDKKNSV 60
Qy 61 CDSSLDIATQLLSNVKLTITSESEKNSLKNKWSITCKSENAETRKVELVPKDPKDL 120
Db 61 CDSSLDIATQLLSNVKLTITSESEKNSLKNKWSITCKSENAETRKVELVPKDPKDL 120
Qy 121 VLAPLDHVNDFTWYKNRKKSIGIKPVKNKIGLSIIPTFNRSRILDTLACLNVQKTY 180
Db 121 VLAPLDHVNDFTWYKNRKKSIGIKPVKNKIGLSIIPTFNRSRILDTLACLNVQKTY 180
Qy 181 PFEVVVADGSKENLLTVQKYEQLDITKYVRQKDYQOLCAVRNLGURTAKYDFVSILD 240
Db 181 PFEVVVADGSKENLLTVQKYEQLDITKYVRQKDYQOLCAVRNLGURTAKYDFVSILD 240
Qy 241 CDMAQQQLWVHSYLTLELLEDNDIVLIGPRKYVDTHNITAEQFLNDPILIESPETATNN 300
Db 241 CDMAQQQLWVHSYLTLELLEDNDIVLIGPRKYVDTHNITAEQFLNDPILIESPETATNN 300
Qy 301 PSITSGNISLDWRLEHFKKTDNLRCDSPFRFYFVAGNVAFSEKWLNVKGVDFEENHWG 360
Db 301 PSITSGNISLDWRLEHFKKTDNLRCDSPFRFYFVAGNVAFSEKWLNVKGVDFEENHWG 360
Qy 361 GEDVEFGYRLFAKGCFFRVYDGGMAIHQEPGKNETEREAGKSTLTKIVKEKVPYIRK 420
Db 361 GEDVEFGYRLFAKGCFFRVYDGGMAIHQEPGKNETEREAGKSTLTKIVKEKVPYIRK 420
Qy 421 LLPEDSHHRIPLYSIYIPAYNCANYTORCVDSALNQTVDVLEVCINDGSTDNTLEVI 480
Db 421 LLPEDSHHRIPLYSIYIPAYNCANYTORCVDSALNQTVDVLEVCINDGSTDNTLEVI 480
Qy 481 NKLYGNNPRVIRMSKPNNGGSIASASNAVSFAKYYIGQLSDSDYLEPDAVELCKEFLKD 540

Db 481 NKLYGNNPRVIRMSKPNNGGSIASASNAVSFAKYYIGQLSDSDYLEPDAVELCKEFLKD 540
Qy 541 KTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTMTAMIAHFRMFTIRAWHLTDGFEN 600
Db 541 KTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTMTAMIAHFRMFTIRAWHLTDGFEN 600
Qy 601 IENAVDYDMFLKLSEVGKFKHLNKCYNRVLHGDNTSIKKLGIOKKNHFFVNVNOSLNROG 660
Db 601 IENAVDYDMFLKLSEVGKFKHLNKCYNRVLHGDNTSIKKLGIOKKNHFFVNVNOSLNROG 660
Qy 661 INYNYDKFDDDESRYIFNKTAETAYOEEMDLKDLKLNQNDKAKIAVSIFYPTNLGLV 720
Db 661 INYNYDKFDDDESRYIFNKTAETAYOEEMDLKDLKLNQNDKAKIAVSIFYPTNLGLV 720
Qy 721 KKLNNIIEYNKNIFVIIHVDKNHLTPDIKKEILAFYHKKHQNILLNNDISYTSNRLIK 780
Db 721 KKLNNIIEYNKNIFVIIHVDKNHLTPDIKKEILAFYHKKHQNILLNNDISYTSNRLIK 780
Qy 781 TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSALTDHWTIEKIN 840
Db 781 TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSALTDHWTIEKIN 840
Qy 841 AHPPEKLIKTYFNDNLSNMVKGASQGMFKYALPHELLTIIKEVITSCQIDSVPY 900
Db 841 AHPPEKLIKTYFNDNLSNMVKGASQGMFKYALPHELLTIIKEVITSCQIDSVPY 900
Qy 901 NTEDIWFQFALLILEKKTGHVFNKSTSTLTYPWPKLQWNTNEQIOSAKKGENIPVNFII 960
Db 901 NTEDIWFQFALLILEKKTGHVFNKSTSTLTYPWPKLQWNTNEQIOSAKKGENIPVNFII 960
Qy 961 NSITL 965
Db 961 NSITL 965

RESULT 4
PCT-US99-07289-3
; Sequence 3, Application PC/TUS9907289
; GENERAL INFORMATION:
; APPLICANT: Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: NUCLEIC ACID ENCODING HYALURONAN SYNTHASE AND METHODS
; TITLE OF INVENTION: OF USE
; FILE REFERENCE: 617022-7
; CURRENT APPLICATION NUMBER: PCT/US99/07289
; CURRENT FILING DATE: 1999-04-01
; EARLIER APPLICATION NUMBER: 60/080,414
; EARLIER FILING DATE: 1998-04-02
; EARLIER APPLICATION NUMBER: 60/178,851
; EARLIER FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Pasteurella multocida
PCT-US99-07289-3

Query Match 98.6%; Score 5017.5; DB 1; Length 972;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 953; Conservative 5; Mismatches 7; Indels 7; Gaps 1;
Qy 1 MNTLSQAIKAYNSNDYELALKLFKSAETGRKIVFQIIKCKEKLST-----NSYVS 53
Db 1 MNTLSQAIKAYNCNDYELALKLFKSAETGRKIVFQIIKCKEKLSTNSYVSEDSYVS 60
Qy 54 EDKNSVCDSSLDIATQLLSNVKLTITSESEKNSLKNKWSITCKSENAETRKVELVP 113
Db 61 EDKNSVCDSSLDIATQLLSNVKLTITSESEKNSLKNKWSITCKSENAETRKVELVP 120
Qy 114 KDFPKDLVLAPLDHVNDFTWYKNRKKSIGIKPVKNKIGLSIIPTFNRSRILDTLACL 173
Db 121 KDFPKDLVLAPLDHVNDFTWYKNRKKSIGIKPVKNKIGLSIIPTFNRSRILDTLACL 180

SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Pasteurella multocida (PmCS)
PCT-US99-07289-8

Query Match 98.6%; Score 5017.5; DB 1; Length 972;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 953; Conservative 5; Mismatches 7; Indels 7; Gaps 1;

QY 1 MNTLSQAIKAYNSNDYELALKLFPEKSAETVGRKIVVEQIIKCKEKLST-----NSVVS 53
Db 1 MNTLSQAIKAYNCNDYELALKLFPEKSAETVGRKIVVEQIIKCKEKLSTNSVSDNSVS 60
QY 54 EDKNSVCDSDSLDTATOLLNSNVKLLTSESEKNSLKNKWSITGKKSSENAEIRKVELVP 113
Db 61 EDKNSVCDSDSLDTATOLLNSNVKLLTSESEKNSLKNKWSITGKKSSENAEIRKVELVP 120
QY 114 KDFPKDLVLAFLPDHVNDFTWYKNRKSLSGKIPVKNKNIIGLSIIITPTNRSRILDTIACL 173
Db 121 KDFPKDLVLAFLPDHVNDFTWYKNRKSLSGKIPVKNKNIIGLSIIITPTNRSRILDTIACL 180
QY 174 VNQKTNTPFEVWVADDSKENLLTIIVQYEQKLDIKYVROKDYGVOLCAVRNLGLRTAKY 233
Db 181 VNQKTNTPFEVWVADDSKENLLTIIVQYEQKLDIKYVROKDYGVOLCAVRNLGLRTAKY 240
QY 234 DFVSLDCDMPAQLWVHSYLTLELNDIVLIGPRKYVDTHNTAEQFLNDPYLIESLP 293
Db 241 DFVSLDCDMPAQLWVHSYLTLELNDIVLIGPRKYVDTHNTAEQFLNDPYLIESLP 300
QY 294 ETATNNPSTISKGNISLDWRLEHFKTDNLRLCDSPPRYFVAGNVAFSKWLNKVGWFD 353
Db 301 ETATNNPSTISKGNISLDWRLEHFKTDNLRLCDSPPRYFVAGNVAFSKWLNKVGWFD 360
QY 354 EEFNHGGEDEVFGYRLFAKGCFFRVIDGMAIHQEPGKENETREAGSKITLKIIVKEK 413
Db 361 EEFNHGGEDEVFGYRLFAKGCFFRVIDGMAIHQEPGKENETREAGSKITLKIIVKEK 420
QY 414 VPYIRKLLPIEDSHIHRIPVLSIIPAYNCANYTORCVDSALNOTVVDLEVCIKNDGST 473
Db 421 VPYIRKLLPIEDSHIHRIPVLSIIPAYNCANYTORCVDSALNOTVVDLEVCIKNDGST 480
QY 474 DNTLEVINKLGNPRVRIMSKPNNGSIASASNAAYSFAGYIIGQLSDSDYVLEPAVELC 533
Db 481 DNTLEVINKLGNPRVRIMSKPNNGSIASASNAAYSFAGYIIGQLSDSDYVLEPAVELC 540
QY 534 LKEFLKDKTFLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 593
Db 541 LKEFLKDKTFLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 600
QY 594 TDGFNENIENAVDYDMFLKSEVGKFKHLNKICYNRVLHGDNTSIKKLGIOKKNHFVVYN 653
Db 601 TDGFNENIENAVDYDMFLKSEVGKFKHLNKICYNRVLHGDNTSIKKLGIOKKNHFVVYN 660
QY 654 QSLNRQGINYNVDKFDLDESRYIFNKTAEOEEMDMKDLKLIQNKDKAKIAVSIFYP 713
Db 661 QSLNRQGINYNVDKFDLDESRYIFNKTAEOEEMDMKDLKLIQNKDKAKIAVSIFYP 720
QY 714 NTLNGLVKKLNIIIEYNKNIIFVILHVDKNHLPDIIKEILAFYHKHQNILLNNDISYY 773
Db 721 NTLNGLVKKLNIIIEYNKNIIFVILHVDKNHLPDIIKEILAFYHKHQNILLNNDISYY 780
QY 774 TSNRLIKTEAHLNSNKLSQLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMFNSALTH 833
Db 781 TSNRLIKTEAHLNSNKLSQLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMFNSALTH 840
QY 834 DWIEKINAHPPPKKLIKTYFNNDLRSNMVKGASQGMFKYALPHELLTIIKEVITSQS 893
Db 841 DWIEKINAHPPPKKLIKTYFNNDLRSNMVKGASQGMFKYALPHELLTIIKEVITSQS 900
QY 894 IDSVPEYNTEDIWFQFALLILEKKTGHVFNKTTSTLTYPWERKQWLTNEQIOSAKKGENI 953
Db 954 IDSVPEYNTEDIWFQFALLILEKKTGHVFNKTTSTLTYPWERKQWLTNEQIOSAKKGENI 960

QY 174 VNQKTNTPFEVWVADDSKENLLTIIVQYEQKLDIKYVROKDYGVOLCAVRNLGLRTAKY 233
Db 181 VNQKTNTPFEVWVADDSKENLLTIIVQYEQKLDIKYVROKDYGVOLCAVRNLGLRTAKY 240
QY 234 DFVSLDCDMPAQLWVHSYLTLELNDIVLIGPRKYVDTHNTAEQFLNDPYLIESLP 293
Db 241 DFVSLDCDMPAQLWVHSYLTLELNDIVLIGPRKYVDTHNTAEQFLNDPYLIESLP 300
QY 294 ETATNNPSTISKGNISLDWRLEHFKTDNLRLCDSPPRYFVAGNVAFSKWLNKVGWFD 353
Db 301 ETATNNPSTISKGNISLDWRLEHFKTDNLRLCDSPPRYFVAGNVAFSKWLNKVGWFD 360
QY 354 EEFNHGGEDEVFGYRLFAKGCFFRVIDGMAIHQEPGKENETREAGSKITLKIIVKEK 413
Db 361 EEFNHGGEDEVFGYRLFAKGCFFRVIDGMAIHQEPGKENETREAGSKITLKIIVKEK 420
QY 414 VPYIRKLLPIEDSHIHRIPVLSIIPAYNCANYTORCVDSALNOTVVDLEVCIKNDGST 473
Db 421 VPYIRKLLPIEDSHIHRIPVLSIIPAYNCANYTORCVDSALNOTVVDLEVCIKNDGST 480
QY 474 DNTLEVINKLGNPRVRIMSKPNNGSIASASNAAYSFAGYIIGQLSDSDYVLEPAVELC 533
Db 481 DNTLEVINKLGNPRVRIMSKPNNGSIASASNAAYSFAGYIIGQLSDSDYVLEPAVELC 540
QY 534 LKEFLKDKTFLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 593
Db 541 LKEFLKDKTFLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 600
QY 594 TDGFNENIENAVDYDMFLKSEVGKFKHLNKICYNRVLHGDNTSIKKLGIOKKNHFVVYN 653
Db 601 TDGFNENIENAVDYDMFLKSEVGKFKHLNKICYNRVLHGDNTSIKKLGIOKKNHFVVYN 660
QY 654 QSLNRQGINYNVDKFDLDESRYIFNKTAEOEEMDMKDLKLIQNKDKAKIAVSIFYP 713
Db 661 QSLNRQGINYNVDKFDLDESRYIFNKTAEOEEMDMKDLKLIQNKDKAKIAVSIFYP 720
QY 714 NTLNGLVKKLNIIIEYNKNIIFVILHVDKNHLPDIIKEILAFYHKHQNILLNNDISYY 773
Db 721 NTLNGLVKKLNIIIEYNKNIIFVILHVDKNHLPDIIKEILAFYHKHQNILLNNDISYY 780
QY 774 TSNRLIKTEAHLNSNKLSQLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMFNSALTH 833
Db 781 TSNRLIKTEAHLNSNKLSQLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMFNSALTH 840
QY 834 DWIEKINAHPPPKKLIKTYFNNDLRSNMVKGASQGMFKYALPHELLTIIKEVITSQS 893
Db 841 DWIEKINAHPPPKKLIKTYFNNDLRSNMVKGASQGMFKYALPHELLTIIKEVITSQS 900
QY 894 IDSVPEYNTEDIWFQFALLILEKKTGHVFNKTTSTLTYPWERKQWLTNEQIOSAKKGENI 953
Db 901 IDSVPEYNTEDIWFQFALLILEKKTGHVFNKTTSTLTYPWERKQWLTNEQIOSAKKGENI 960
QY 954 PVNKFIIINSITL 965
Db 961 PVNKFIIINSITL 972

RESULT 5
PCT-US99-07289-8
; Sequence 8, Application PC/TUS9907289
; GENERAL INFORMATION:
; APPLICANT: Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: NUCLEIC ACID ENCODING HYALURONAN SYNTHASE AND METHODS
; TITLE OF INVENTION: OF USE
; FILE REFERENCE: 617022-7
; CURRENT APPLICATION NUMBER: PCT/US99/07289
; CURRENT FILING DATE: 1999-04-01
; EARLIER APPLICATION NUMBER: 60/080,414
; EARLIER FILING DATE: 1998-04-02
; EARLIER APPLICATION NUMBER: 60/178,851
; EARLIER FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 29

Db 901 IDSVPEYNTEDIWFQFALLILEKKTGHVFNKTSTLTYPWPWKQLQWNTNEQIOSAKKGENI 960
Qy 954 PVNKFIINSITL 965
Db 961 PVNKFIINSITL 972
RESULT 6
US-09-283-402-3
; Sequence 3, Application US/09283402
; GENERAL INFORMATION:
; APPLICANT: Paul DeAngelis
; TITLE OF INVENTION: DNA ENCODING HYALURONAN SYNTHASE FROM PASTURELLA
; FILE REFERENCE: 5820.530
; CURRENT APPLICATION NUMBER: US/09/283.402
; CURRENT FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-09-283-402-3
Query Match 98.6%; Score 5017.5; DB 16; Length 972;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 953; Conservative 5; Mismatches 7; Indels 7; Gaps 1;
Qy 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGRIKIVFEQIIKCKEKLST-----NSVYS 53
Db 1 MNTLSQAIKAYNCNDYELALKLFKSAETYGRIKIVFEQIIKCKEKLSTNSVYSEDNSVYS 60
Qy 54 EDKKNVCDSSLDIATOLLISNVKLTLSSEKNSLKNKWSITGKKSNAEIRKVELVP 113
Db 61 EDKKNVCDSSLDIATOLLISNVKLTLSSEKNSLKNKWSITGKKSNAEIRKVELVP 120
Qy 114 KDFPKDLVLA PLDPHVDFTWYKRRKSLGIPKVNKNIGLSIIPTFNRSRILDTITLACL 173
Db 121 KDFPKDLVLA PLDPHVDFTWYKRRKRLGIPKVNKNIGLSIIPTFNRSRILDTITLACL 180
Qy 174 VNQKTNPFVAVVADGSGKENLLTIQVKEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKY 233
Db 181 VNQKTNPFVAVVADGSGKENLLTIQVKEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKY 240
Qy 234 DFVSILDCDMPAQOLWVHSYLTLEDDNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 293
Db 241 DFVSILDCDMPAQOLWVHSYLTLEDDNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 300
Qy 294 ETATNNNPSITSGNISLDWRLEHFKKTDLNRLCDSPPRYFVAGNVAFSKENLKNVGVWF 353
Db 301 ETATNNNPSITSGNISLDWRLEHFKKTDLNRLCDSPPRYFVAGNVAFSKENLKNVGVWF 360
Qy 354 EEFNHWGGEDVEFGYRLFAKGCFFRIDGMAIHOEPGPKENETEREAGSKITLKIIVEK 413
Db 361 EEFNHWGGEDVEFGYRLFAKGCFFRIDGMAIHOEPGPKENETEREAGSKITLKIIVEK 420
Qy 414 VPYIYRKLPTEDSHIRIPLVSIYIPAYNCANYIQRCDVSALAQTVVDVLEVCINCDSGT 473
Db 421 VPYIYRKLPTEDSHIRIPLVSIYIPAYNCANYIQRCDVSALAQTVVDVLEVCINCDSGT 480
Qy 474 DNTLEVINKLYGNPRVIRIMSKPNGGIASANAASVFAKGYIIGQLSDDDYLEPDAVELC 533
Db 481 DNTLEVINKLYGNPRVIRIMSKPNGGIASANAASVFAKGYIIGQLSDDDYLEPDAVELC 540
Qy 534 LKEFLKDKTTLACVYTTNRNVPDGLSANGYNWPEFSREKLTUTAMIAHFFRMFTIRAWHL 593
Db 541 LKEFLKDKTTLACVYTTNRNVPDGLSANGYNWPEFSREKLTUTAMIAHFFRMFTIRAWHL 600
Qy 594 TDGFNENIENAVDYDMFLKLEVGKFKHLNKCICYNRVLHGDNSTIKKIGIOKKNHFVVVN 653
Db 601 TDGFNENIENAVDYDMFLKLEVGKFKHLNKCICYNRVLHGDNSTIKKIGIOKKNHFVVVN 660

Qy 654 QSLNRQGINYNYDKFDDDESRYKIFNKTAETAEQEMDMKDLKLIQNKDAKIAVSIFYP 713
Db 661 QSLNRQGINYNYDKFDDDESRYKIFNKTAETAEQEMDMKDLKLIQNKDAKIAVSIFYP 720
Qy 714 NTLNGLVKKLNIIIEYNKNIFVIIHVDKNHILTPDIKKELAFYHKHQNILLNNDISYY 773
Db 721 NTLNGLVKKLNIIIEYNKNIFVIIHLDKNHILTPDIKKELAFYHKHQNILLNNDISYY 780
Qy 774 TSNRLIKTEAHLNKLNSQLNLCNEYIIFDNHDSLFVKNDSYAYMKKYDVGMFNSALTH 833
Db 781 TSNRLIKTEAHLNKLNSQLNLCNEYIIFDNHDSLFVKNDSYAYMKKYDVGMFNSALTH 840
Qy 834 DWIEKINAHPPPKLIKTYFNDNDRSMNVKGASQGMFMKYALPHELHLLTIKEVITSCOS 893
Db 841 DWIEKINAHPPPKLIKTYFNDNDRSMNVKGASQGMFMKYALRHALLTIKEVITSCOS 900
Qy 894 IDSVPEYNTEDIWFQFALLILEKKTGHVFNKTSTLTYPWPWKQLQWNTNEQIOSAKKGENI 953
Db 901 IDSVPEYNTEDIWFQFALLILEKKTGHVFNKTSTLTYPWPWKQLQWNTNEQIOSAKKGENI 960
Qy 954 PVNKFIINSITL 965
Db 961 PVNKFIINSITL 972
RESULT 7
US-09-283-402-8
; Sequence 8, Application US/09283402
; GENERAL INFORMATION:
; APPLICANT: Paul DeAngelis
; TITLE OF INVENTION: DNA ENCODING HYALURONAN SYNTHASE FROM PASTURELLA
; FILE REFERENCE: 5820.530
; CURRENT APPLICATION NUMBER: US/09/283.402
; CURRENT FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Pasteurella multocida (PmCS)
US-09-283-402-8
Query Match 98.6%; Score 5017.5; DB 16; Length 972;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 953; Conservative 5; Mismatches 7; Indels 7; Gaps 1;
Qy 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGRIKIVFEQIIKCKEKLST-----NSVYS 53
Db 1 MNTLSQAIKAYNCNDYELALKLFKSAETYGRIKIVFEQIIKCKEKLSTNSVYSEDNSVYS 60
Qy 54 EDKKNVCDSSLDIATOLLISNVKLTLSSEKNSLKNKWSITGKKSNAEIRKVELVP 113
Db 61 EDKKNVCDSSLDIATOLLISNVKLTLSSEKNSLKNKWSITGKKSNAEIRKVELVP 120
Qy 114 KDFPKDLVLA PLDPHVDFTWYKRRKSLGIPKVNKNIGLSIIPTFNRSRILDTITLACL 173
Db 121 KDFPKDLVLA PLDPHVDFTWYKRRKRLGIPKVNKNIGLSIIPTFNRSRILDTITLACL 180
Qy 174 VNQKTNPFVAVVADGSGKENLLTIQVKEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKY 233
Db 181 VNQKTNPFVAVVADGSGKENLLTIQVKEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKY 240
Qy 234 DFVSILDCDMPAQOLWVHSYLTLEDDNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 293
Db 241 DFVSILDCDMPAQOLWVHSYLTLEDDNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 300
Qy 294 ETATNNNPSITSGNISLDWRLEHFKKTDLNRLCDSPPRYFVAGNVAFSKENLKNVGVWF 353
Db 301 ETATNNNPSITSGNISLDWRLEHFKKTDLNRLCDSPPRYFVAGNVAFSKENLKNVGVWF 360
Qy 354 EEFNHWGGEDVEFGYRLFAKGCFFRIDGMAIHOEPGPKENETEREAGSKITLKIIVEK 413
Db 354 EEFNHWGGEDVEFGYRLFAKGCFFRIDGMAIHOEPGPKENETEREAGSKITLKIIVEK 413

Db	361	EEFNHGGEDVEFGYRLFPKGCFFRVIDGGMAYHQPPGKENETEREAGKSITLKIVKEK	420
Qy	414	VPYIYKLLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVVLEVCICNDGST	473
Db	421	VPYIYKLLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVVLEVCICNDGST	480
Qy	474	DNTLEVINLKYNNPRVIRMSKPGNGGASASNAAFSAKGYIIGQLSDSDYLEPDAVELC	533
Db	481	DNTIEVINLKYNNPRVIRMSKPGNGGASASNAAFSAKGYIIGQLSDSDYVLEPDAVELC	540
Qy	534	LKEFLDKDTLACVYTTNRNVPDGLSIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL	593
Db	541	LKEFLDKDTLACVYTTNRNVPDGLSIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL	600
Qy	594	TGDFNENIENAVDYMFLKSEVGKFKHLNKICYNRVLHGDNTSIKKGIOKKNHFVVVN	653
Db	601	TGDFNENIENAVDYMFLKSEVGKFKHLNKICYNRVLHGDNTSIKKGIOKKNHFVVVN	660
Qy	654	QSLNRQGINYYNDKFDLDESRYIFNKTAEYQEEEDMLDKLQIONKDAKIAVSIFYP	713
Db	661	QSLNRQGINYYNDKFDLDESRYIFNKTAEYQEEEDMLDKLQIONKDAKIAVSIFYP	720
Qy	714	NTNLGLVKLLNNIIEYNKNIFVILHLDKNHLTPDIKKEILAFYHKHQVNILLNNDISY	773
Db	721	NTNLGLVKLLNNIIEYNKNIFVILHLDKNHLTPDIKKEILAFYHKHQVNILLNNDISY	780
Qy	774	TSNRLIKTEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSALTH	833
Db	781	TSNRLIKTEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSALTH	840
Qy	834	DWIEKINAHPPFKLIKTYFNDNDRSMNVKASQGMFKYALPHELLTIIKEVITSCQS	893
Db	841	DWIEKINAHPPFKLIKTYFNDNDRSMNVKASQGMFKYALPHELLTIIKEVITSCQS	900
Qy	894	IDSVPYNTEDIWFQALLILEKKTGHVNKTSTLTYPWERKLOWTNEQIOSAKKGENI	953
Db	901	IDSVPYNTEDIWFQALLILEKKTGHVNKTSTLTYPWERKLOWTNEQIOSAKKGENI	960
Qy	954	PVNKFIINSITL 965	
Db	961	PVNKFIINSITL 972	
RESULT 8			
US-09-469-200-3			
; Sequence 3, Application US/09469200			
; GENERAL INFORMATION:			
; APPLICANT: Board of Regents of the University of Oklahoma			
; TITLE OF INVENTION: NUCLEIC ACID ENCODING HYALURONAN SYNTHASE AND METHODS			
; TITLE OF INVENTION: OF USE			
; FILE REFERENCE: 617022-7			
; CURRENT APPLICATION NUMBER: US/09/469,200			
; PRIOR FILING DATE: 1999-12-21			
; PRIOR APPLICATION NUMBER: 60/080,414			
; PRIOR FILING DATE: 1998-04-02			
; PRIOR APPLICATION NUMBER: 60/178,851			
; PRIOR FILING DATE: 1998-10-26			
; NUMBER OF SEQ ID NOS: 29			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 3			
; LENGTH: 972			
; TYPE: PRT			
; ORGANISM: Pasteurella multocida			
US-09-469-200-3			
Query Match 98.6%; Score 5017.5; DB 18; Length 972;			
Best Local Similarity 98.0%; Pred. No. 0;			
Matches 953; Conservative 5; Mismatches 7; Indels 7; Gaps 1;			
Qy	1	MNTLSQAIKAYNSNDYELAKLFKESAEITYGRKIVEFQIICKKEKIST-----NSYVS	53
Db	1	MNTLSQAIKAYNCNDYELAKLFKESAEITYGRKIVEFQIICKKEKISTNSYVSEDNSYVS	60

RESULT 9
US-09-469-200-8
; Sequence 8, Application US/09469200
; GENERAL INFORMATION:
; APPLICANT: Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: NUCLEIC ACID ENCODING HYALURONAN SYNTHASE AND METHODS
; TITLE OF INVENTION: OF USE
; FILE REFERENCE: 617022-7

```

; CURRENT APPLICATION NUMBER: US/09/469,200
; CURRENT FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: 60/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 8
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Pasteurella multocida (PmCs)
; US-09-469-200-8

Query Match      98.6%; Score 5017.5; DB 18; Length 972;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 953; Conservative 5; Mismatches 7; Indels 7; Gaps 1;

Qy 1 MNTLSQAIKAYNSNDYELALKLFKSAETGRKIVFQIIKCKEKLST-----NSYVS 53
Db 1 MNTLSQAIKAYNSNDYELALKLFKSAETGRKIVFQIIKCKEKLSTNSYVS 60

Qy 54 EDKNSVCDSSLDIATQLLSNVKKLTLSSEKSLKNKWSITGKKSNAEIRKVELVP 113
Db 61 EDKNSVCDSSLDIATQLLSNVKKLTLSSEKSLKNKWSITGKKSNAEIRKVELVP 120

Qy 114 KDFPKDLVLPDHPVNDFTWYKNRKKSIGIKPVKNKTGLSTIIPFNRSRILDTLACL 173
Db 121 KDFPKDLVLPDHPVNDFTWYKNRKKSIGIKPVKNKTGLSTIIPFNRSRILDTLACL 180

Qy 174 VNQNTNPFVYVADGSGKENLLTTVQYEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKY 233
Db 181 VNQNTNPFVYVADGSGKENLLTTVQYEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKY 240

Qy 234 DFVSLDCDMPAQQLWVHSYLTLELNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 293
Db 241 DFVSLDCDMPAQQLWVHSYLTLELNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 300

Qy 294 ETATNNNFSITSGNISLDWRLHFKTDNLRCLDPSPRYFVAGNVAFSKWLNKVGWFD 353
Db 301 ETATNNNFSITSGNISLDWRLHFKTDNLRCLDPSPRYFVAGNVAFSKWLNKVGWFD 360

Qy 354 EEFNWGGEDVEFGYRLFAKGFRRVIDGGMIAHOEPCKENETEREAGKSITLKIIVREK 413
Db 361 EEFNWGGEDVEFGYRLFAKGFRRVIDGGMIAHOEPCKENETEREAGKSITLKIIVREK 420

Qy 414 VPIYIRKLLPTEDSHIIRIPLVSIYPAYNCANYTORCVDSALNQTVDVLEVCINDGST 473
Db 421 VPIYIRKLLPTEDSHIIRIPLVSIYPAYNCANYTORCVDSALNQTVDVLEVCINDGST 480

Qy 474 DNTLEVINKLYGNPRVRIMSKPNGGSIASASNAVSFAKGYIGQLDSDDYLEPDAVELC 533
Db 481 DNTLEVINKLYGNPRVRIMSKPNGGSIASASNAVSFAKGYIGQLDSDDYLEPDAVELC 540

Qy 534 LKEFLKDKTLACVYTTNRNPNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 593
Db 541 LKEFLKDKTLACVYTTNRNPNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 600

Qy 594 TDGFNENENAVDYMFLKLSVGFKHLNKICYNRVLHGDNTSKKLGIOQKKNHFVVYN 653
Db 601 TDGFNENENAVDYMFLKLSVGFKHLNKICYNRVLHGDNTSKKLGIOQKKNHFVVYN 660

Qy 654 QSLNRQGINNYNDKFDLDESRRKIFNKTAQYQEMDMKDLKLTIONKDAKIAVSIFYP 713
Db 661 QSLNRQGINNYNDKFDLDESRRKIFNKTAQYQEMDMKDLKLTIONKDAKIAVSIFYP 720

Qy 714 NTLNGLVKKLNNIIEYNKNFVIIHLVDKNHLLTPDIKKEILAFYKHQVNNLLNNDISY 773
Db 721 NTLNGLVKKLNNIIEYNKNFVIIHLVDKNHLLTPDIKKEILAFYKHQVNNLLNNDISY 780

Qy 774 TSNRLIKTEAHLNNTKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMNFSALTH 833
Db 781 TSNRLIKTEAHLNNTKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMNFSALTH 840
```

```

Qy 834 DWIEKINAHPPFKKLIKTYFNDNLRSMNVKGASQGMKMYALPHELLTIIKEVITSQS 893
Db 841 DWIEKINAHPPFKKLIKTYFNDNLRSMNVKGASQGMKMYALPHELLTIIKEVITSQS 900

Qy 894 IDSVPYENTEDIWFQFALLILEKKTGHVFNKTSLTLYMPWERKLOMTNEQIOSAKKGENI 953
Db 901 IDSVPYENTEDIWFQFALLILEKKTGHVFNKTSLTLYMPWERKLOMTNEQIOSAKKGENI 960

Qy 954 PVNKFIIINSITL 965
Db 961 PVNKFIIINSITL 972

RESULT 10
PCT-US99-07289-9
; Sequence 9, Application PC/TUS9907289
; GENERAL INFORMATION:
; APPLICANT: Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: NUCLEIC ACID ENCODING HYALURONAN SYNTHASE AND METHODS
; TITLE OF INVENTION: OF USE
; FILE REFERENCE: 617022-7
; CURRENT APPLICATION NUMBER: PCT/US99/07289
; CURRENT FILING DATE: 1999-04-01
; EARLIER APPLICATION NUMBER: 60/080,414
; EARLIER FILING DATE: 1998-04-02
; EARLIER APPLICATION NUMBER: 60/178,851
; EARLIER FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 9
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (17)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (62)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (90)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (113)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (158)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (164)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (192)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (201)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (206)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
```

NAME/KEY: MOD_RES
LOCATION: (211)
OTHER INFORMATION: either Asn, Asp, Glu or Gln
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (225)
OTHER INFORMATION: either Phe or Tyr
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (233)
OTHER INFORMATION: either Met or Leu
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (243)
OTHER INFORMATION: either Ile or Val
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (253)
OTHER INFORMATION: either Asn, Asp, Glu or Gln
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (279)
OTHER INFORMATION: either Ile or Val
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (288)
OTHER INFORMATION: either Asn, Asp, Glu or Gln
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (292)
OTHER INFORMATION: either Asn, Asp, Glu or Gln
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (316)
OTHER INFORMATION: either Ile or Val
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (329)
OTHER INFORMATION: either Phe or Tyr
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (340)
OTHER INFORMATION: either Phe or Tyr
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (405)
OTHER INFORMATION: either Asn, Asp, Glu or Gln
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (439)
OTHER INFORMATION: either Ile or Val
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (744)
OTHER INFORMATION: either Ile or Val
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (952)
OTHER INFORMATION: either Asn, Asp, Glu or Gln

Query Match 90.4%; Score 4601.5; DB 1; Length 972;
Best Local Similarity 89.9%; Pred. No. 0;
Matches 874; Conservative 28; Mismatches 63; Indels 7; Gaps 1;
QY 1 MNTLSQAIKAYNSNDYELAKLIFKSAEYGRKIVEFQIIKCKELSTNS-----YYS 53
Db 1 MNTLSQAIKAYNCNDYXLKLFKSAEYGRKIVEFQIIKCKELKUSANPSVNEAMLSYN 60
QY 54 EDKNSVCSSLDIATOLLISNVKLTLSSEKNSLKNKWSKITGKKSNAEIRKVELYP 113
Db 61 EXEKNVCSPLDIATOLLISNVKLTLSSEKNSLKNKWLITERKKSNAEIRKVELYP 120

QY 114 KDFPKDLVLAPDPHVNDEFTWYNRKKSLGKIPVKNKIGLSIIIPTFNRSRILDTIACL 173
Db 121 KDFPKDLVLAPDPHVNDEFTWYNRKKRLGKIPKNGXGLSIIIXPTFNRRAPIITLACL 180
QY 174 VNOKTYPFEVVVADDSKKNLLTIYOKYEQKLDIKVYRQKDYQYOLCAVNLGLRTAKY 233
Db 181 VNOKTYPFEVVVADDSKEXLLPIXROYEXKLDIRVYRQKNGXQACAARNXGLRLAKY 240
QY 234 DFVSLDCDMAPQOLWVHSYLTLELDDNDIVLIGPRKYVDTHNTAEQFLNDPLYIESLP 293
Db 241 DFVSLDCDMAPQOLWVHSYLAELLEDDDTIIGPRKYDTONDAEFLXASLIESLP 300
QY 294 ETATNNPSTISKGNISLDMRLEHFKTDNRLCDSFPRYFVAGNVAFSKWLAKVGHFD 353
Db 301 ETATNNPAAKGEGNXSLDMRLEQFEKTXNRLCDSFPRFAAGNVAFKWLAKNGSGFD 360
QY 354 EEFNHGGEDEVFGYRLFAKCCFVRIDGGMAIHOEPGKENEREAGKSTIKLIVREK 413
Db 361 EEFNHGGEDEVFGYRLFRKGCFFERTIDGGMAIHOEPGKENEREAGKSTIKLIVREK 420
QY 414 VPIYRKLPTEDSHIHRIPLVSIYIPAYNCANYIQRVDSALNQTVVDLEVYCINDGST 473
Db 421 VPIYRKLPTEDSHINRXPVSIYIPAYNCANYIQRVDSALNQTVVDLEVYCINDGST 480
QY 474 DNTLEVINKLYGNPRVRIMSKPNGGSIASANAASVFAKYYIGOLDSDDYLEPAVELC 533
Db 481 DNTLEVINKLYGNPRVRIMSKPNGGSIASANAASVFAKYYIGOLDSDDYLEPAVELC 540
QY 534 LKEFLDKTLACVYTTNENVPDGLIANGYNWPEFSREKLTMTAMIAHFRMFTIRAWHL 593
Db 541 LKEFLDKTLACVYTTNENVPDGLIANGYNWPEFSREKLTMTAMIAHFRMFTIRAWHL 600
QY 594 TDGFENIENAVDYDMFLKSEVGKFKHLNKICYNRVLHGDNTSIKKLGIOKKNHFVVYN 653
Db 601 TDGFENIENAVDYDMFLKSEVGKFKHLNKICYNRVLHGDNTSIKKLGIOKKNHFVVYN 660
QY 654 QSLNRQGINYNKYDFDLDESRRKIFNKTAEYOEMDKLKLIONKDAKIAVSIFYP 713
Db 661 QSLNRQGINYNKYDFDLDESRRKIFNKTAEYOEMDKLKLIONKDAKIAVSIFYP 720
QY 714 NTLNGLVKLNNIIEYNKNIFVILHVDKNHLLTPDIKKEILAFYHKHVNILLNNDISY 773
Db 721 NTLNGLVKLNNIIEYNKNIFVILHLDKNHLLTPDIKKEILAFYHKHVNILLNNDISY 780
QY 774 TSNRLIKTEAHLNINKLSQMLNCEYIIFDNDHDSLFVKNDSYAYMKKYDVGMNFSALTH 833
Db 781 TSNRLIKTEAHLNINKLSQMLNCEYIIFDNDHDSLFVKNDSYAYMKKYDVGMNFSALTH 840
QY 834 DWIEKINAHPPPKLLIKTYFENDNDRSMNVKASQGMFKYALPHELLTIIEKIVITSCOS 893
Db 841 DWIEKINAHPPPKLLIKTYFENDNDRSMNVKASQGMFKYALPHELLTIIEKIVITSCOS 900
QY 894 IDSVPYNTEDIWFQFALLILEKKTGHVFNKTSTLTYPMPWERKLTWNTQISAKKGENI 953
Db 901 IDSVPYNTEDIWFQFALLILEKKTGHVFNKTSTLTYPMPWERKLTWNTQISAKKGENI 960
QY 954 PVNKFIIINSITL 965
Db 961 PVNKFIIINSITL 972

RESULT 11
US-09-283-402-9
; Sequence 9, Application US/09283402
; GENERAL INFORMATION:
; APPLICANT: Paul DeAngelis
; TITLE OF INVENTION: DNA ENCODING HYALURONAN SYNTHASE FROM PASTURELLA
; TITLE OF INVENTION: MULTOCIDA AND USES THEREOF
; FILE REFERENCE: 5820.530
; CURRENT APPLICATION NUMBER: US/09/283,402
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0

```
; SEQ ID NO 9
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; NAME/KEY: MOD_RES
; LOCATION: (17)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; NAME/KEY: MOD_RES
; LOCATION: (90)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; NAME/KEY: MOD_RES
; LOCATION: (62)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; NAME/KEY: MOD_RES
; LOCATION: (113)
; OTHER INFORMATION: either Ile or Val
; NAME/KEY: MOD_RES
; LOCATION: (158)
; OTHER INFORMATION: either Ile or Val
; NAME/KEY: MOD_RES
; LOCATION: (164)
; OTHER INFORMATION: either Ile or Val
; NAME/KEY: MOD_RES
; LOCATION: (201)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; NAME/KEY: MOD_RES
; LOCATION: (206)
; OTHER INFORMATION: either Ile or Val
; NAME/KEY: MOD_RES
; LOCATION: (211)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; NAME/KEY: MOD_RES
; LOCATION: (225)
; OTHER INFORMATION: either Phe or Tyr
; NAME/KEY: MOD_RES
; LOCATION: (233)
; OTHER INFORMATION: either Met or Leu
; NAME/KEY: MOD_RES
; LOCATION: (243)
; OTHER INFORMATION: either Ile or Val
; NAME/KEY: MOD_RES
; LOCATION: (253)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; NAME/KEY: MOD_RES
; LOCATION: (279)
; OTHER INFORMATION: either Ile or Val
; NAME/KEY: MOD_RES
; LOCATION: (288)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; NAME/KEY: MOD_RES
; LOCATION: (292)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; NAME/KEY: MOD_RES
; LOCATION: (316)
; OTHER INFORMATION: either Ile or Val
; NAME/KEY: MOD_RES
; LOCATION: (329)
; OTHER INFORMATION: either Phe or Tyr
; NAME/KEY: MOD_RES
; LOCATION: (340)
; OTHER INFORMATION: either Phe or Tyr
; NAME/KEY: MOD_RES
; LOCATION: (405)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; NAME/KEY: MOD_RES
; LOCATION: (435)
; OTHER INFORMATION: either Ile or Val
; NAME/KEY: MOD_RES
```

```
; LOCATION: (744)
; OTHER INFORMATION: either Ile or Val
; NAME/KEY: MOD_RES
; LOCATION: (952)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; US-09-283-402-9

Query Match          90.4%; Score 4601.5; DB 16; Length 972;
Best Local Similarity 89.9%; Pred. No. 0;
Matches 874; Conservative 28; Mismatches 63; Indels 7; Gaps 1;

QY 1 MNTLSQAIKAYNSNDYELAKLFEKSAETVGRKIVFQIKKEKIKSTNS-----YVS 53
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MNTLSQAIKAYNCNDYXLAKLFEKSAETVGRKIVFQIKKEKLSANPSVNEANLSVN 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 54 EDKNSVCDSSLDIATQLLSNVKLTLSSEKNSLKNKWKSTGKSENAERKVELVP 113
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 EXEKNVCDSPLDIATQLLSINVKLTLSXSEKNSLKNKWKLTERRKSENAERAVALVP 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 114 KDFPKDLVLAPLDHVNDFTWYKNRKSLGKIPVKNIGLSIYIPFNRSRIIDITLACL 173
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 KDFPKDLVLAPLDHVNDFTWYKNRKSLGKIPENGXGLSIIIXPTFNRAILDITLACL 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 174 VNOKTNPFEVYVADGSGKENLLTIVQYEQKLDIKYVRQKDYGYQLCAVRNGLRTAKY 233
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 VNOKTNPFEVYVADGSGEXLLPIXRQYEXKLDIRYVRQKNGXQACAAARNXGLRLAKY 240
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 234 DFVSLDCDMAPOQLWVHSYLTLELLEDNDIVLIGPRKYVDTHNITAEOLFNDPYLTESLP 293
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 DFVSLDCDMAPOQLWVHSYLAELLEDNDITIGPRKYXDTQNDIAEXFLNXASLTESLP 300
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 294 ETATNNPITSKGNISLDWRLEHFKKTDNRLCDSPPRFYVAGNVAFSEWLNKVGWFD 353
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 ETATNNPAKGEKNXSLDWRLEQFEKTXNRLCDSPPRFYVAGNVAFSEWLNKSGFFD 360
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 354 EEFNHGGEDVEFGYRLFAKGCFFRVIDGSMATHQPPGKENTETERAGKSTILKIVKEK 413
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 EEFNHGGEDVEFGYRLFRKGCFFRITDGMAYHQBPBGKENETXREAGKNITLDIMREK 420
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 414 VPYIYRKLPIEDSHIRIPLVSIYIPAYNCANYIQRVDSALNQTVLDLEVCINDGST 473
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 VPYIYRKLPIEDSHIRIPLVSIYIPAYNCANYIQRVDSALNQTVLDLEVCINDGST 480
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 474 DNTLEVINKLYGNPRVRIMSKPNGGIASASNAAVSAFAGYIIGQLSDDDYLEPDAYELC 533
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 DNTLEVINKLYGNPRVRIMSKPNGGIASASNAAVSAFAGYIIGQLSDDDYLEPDAYELC 540
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 534 LKEFLKDKTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTMTAMIAHFRMFTIRAWHL 593
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 LKEFLKDKTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTMTAMIAHFRMFTIRAWHL 600
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 594 TDGFENENIENAVDYDMFLKLSVGEKPKHLNKICYNRVLHGDNITSIKKLGIOKKNHFFVVN 653
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 601 TDGFENENIENAVDYDMFLKLSVGEKPKHLNKICYNRVLHGDNITSIKKLGIOKKNHFFVVN 660
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 654 QSLNRQGINYNYDKFDDDESRYIFNKTAEYQEEMDLKDLKIQNKDAKIAVSTFYP 713
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 661 QSLNRQGINYNYDEFDDDESRYIFNKTAEYQEEDILDKIIOKNDKAKIAVSTFYP 720
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 714 NTLNGLVKLLNIIIEYNKNIIFVILHVDKNHLTPDKKEILAFYHKHQVILLNNDISYY 773
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 721 NTLNGLVKLLNIIIEYNKNIIFVILHLDKNHLTPDKKEILAFYHKHQVILLNNDISYY 780
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 774 TSNRLIKTEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGNMFSALTH 833
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 781 TSNRLIKTEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGNMFSALTH 840
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 834 DWIEKINAHPPFKKLIKTYFNDNDRSMNVKGASQGMFKYALPHELLTIIKEVITSCQS 893
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 841 DWIEKINAHPPFKKLIKTYFNDNDRSMNVKGASQGMFKYALAHALLTIIKEVITSCQS 900
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 894 IDSVPEYNTEDIWFQFALLILEKKTGHVFNKSTLTLYMPWPKLQNTNEOIQSAKGENI 953
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 901 IDSVPEYNTEDINWFALLILEKKTGHVFNKSTLTLYMPWPKLQWNTNEQIXSAKRGENI 960

QY 954 PVNKFINSITL 965
|||||

Db 961 PVNKFINSITL 972

RESULT 12

US-09-469-200-9

Sequence 9, Application US/09469200

GENERAL INFORMATION:

APPLICANT: Board of Regents of the University of Oklahoma

TITLE OF INVENTION: NUCLEIC ACID ENCODING HYALURONAN SYNTHASE AND METHODS

TITLE OF INVENTION: OF USE

FILE REFERENCE: 617022-7

CURRENT APPLICATION NUMBER: US/09/469,200

CURRENT FILING DATE: 1999-12-21

PRIOR APPLICATION NUMBER: 60/080,414

PRIOR FILING DATE: 1998-04-02

PRIOR APPLICATION NUMBER: 60/178,851

PRIOR FILING DATE: 1998-10-26

NUMBER OF SEQ ID NOS: 29

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 9

LENGTH: 972

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: consensus

NAME/KEY: MOD_RES

LOCATION: (17)

OTHER INFORMATION: either Asn, Asp, Glu or Gln

NAME/KEY: MOD_RES

LOCATION: (62)

OTHER INFORMATION: either Asn, Asp, Glu or Gln

NAME/KEY: MOD_RES

LOCATION: (90)

OTHER INFORMATION: either Asn, Asp, Glu or Gln

NAME/KEY: MOD_RES

LOCATION: (113)

OTHER INFORMATION: either Ile or Val

NAME/KEY: MOD_RES

LOCATION: (158)

OTHER INFORMATION: either Ile or Val

NAME/KEY: MOD_RES

LOCATION: (164)

OTHER INFORMATION: either Ile or Val

NAME/KEY: MOD_RES

LOCATION: (192)

OTHER INFORMATION: either Ile or Val

NAME/KEY: MOD_RES

LOCATION: (201)

OTHER INFORMATION: either Asn, Asp, Glu or Gln

NAME/KEY: MOD_RES

LOCATION: (206)

OTHER INFORMATION: either Ile or Val

NAME/KEY: MOD_RES

LOCATION: (211)

OTHER INFORMATION: either Asn, Asp, Glu or Gln

NAME/KEY: MOD_RES

LOCATION: (225)

OTHER INFORMATION: either Phe or Tyr

NAME/KEY: MOD_RES

LOCATION: (233)

OTHER INFORMATION: either Met or Leu

NAME/KEY: MOD_RES

LOCATION: (243)

OTHER INFORMATION: either Ile or Val

NAME/KEY: MOD_RES

LOCATION: (253)

OTHER INFORMATION: either Asn, Asp, Glu or Gln

NAME/KEY: MOD_RES

LOCATION: (479)

OTHER INFORMATION: either Ile or Val

NAME/KEY: MOD_RES

LOCATION: (288)

OTHER INFORMATION: either Asn, Asp, Glu or Gln

NAME/KEY: MOD_RES

LOCATION: (292)

OTHER INFORMATION: either Asn, Asp, Glu or Gln

NAME/KEY: MOD_RES

LOCATION: (316)

OTHER INFORMATION: either Ile or Val

NAME/KEY: MOD_RES

LOCATION: (329)

OTHER INFORMATION: either Phe or Tyr

NAME/KEY: MOD_RES

LOCATION: (340)

OTHER INFORMATION: either Phe or Tyr

NAME/KEY: MOD_RES

LOCATION: (405)

OTHER INFORMATION: either Asn, Asp, Glu or Gln

NAME/KEY: MOD_RES

LOCATION: (439)

OTHER INFORMATION: either Ile or Val

NAME/KEY: MOD_RES

LOCATION: (744)

OTHER INFORMATION: either Ile or Val

NAME/KEY: MOD_RES

LOCATION: (952)

OTHER INFORMATION: either Asn, Asp, Glu or Gln

US-09-469-200-9

Query Match 90.4%; Score 4601.5; DB 18; Length 972;

Best Local Similarity 89.9%; Pred. No. 0;

Matches 874; Conservative 28; Mismatches 63; Indels 7; Gaps 1;

QY 1 MNTLSQAIKAYNSNDYELALKLFEKSAETIYGRKIYVEFQIIEKKEKLSN-----YVS 53

Db 1 MNTLSQAIKAYNCNDYXLALKLFEKSAETIYGRKIYVEFQIIEKKEKLSNPNVNEANLSVN 60

QY 54 EDKNSVCDSSLDIATQLLSNVKLTISESEKSLKNKWSITGKKSSNAEIRKVELVP 113

Db 61 EXEKNVCDSPLDIATQLLSNVKLTISXSEKSLKNKWLITEKKSNAEXRAVALVP 120

QY 114 KDFPKDLVLAFLPDHVNDFTWYKNEKKSIGIKPVKNKIGLSIIPFNRSRLDITLACL 173

Db 121 KDFPKDLVLAFLPDHVNDFTWYKNEKKSIGIKPVKNKIGLSIIPFNRSRLDITLACL 180

QY 174 VNQKTNYPFEVVVADDGSGKENLTIYQVKEQKLDIKYVRQKDYGYOLCAVRNLGLRTAKY 233

Db 181 VNQKTNYPFEVVVADDGSGEXLLPIXRQYEXKLDIRYVRQKDXQACAARNKGLRLAKY 240

QY 234 DPFVSLDCDMPQQLWVHSYLTLELLEDNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 293

Db 241 DFXGILDCDMPXQLWVHSYLAELLEDDDDITIGPRKYXDTQNDIDAEFLNXASLIESLP 300

QY 294 ETATNNPSITSGNISLDWRLEHFKKTDNRLCDSPPFRYFVAGNVAFSKEWLNKYGWFD 353

Db 301 ETATNNPAKGEKXSLDWRLEQFEKTXNRLCDSPPFRFAAGNVAFSAKWLNKSGFFD 360

QY 354 EEFNHWGGEDVEFGYRLFAKGCFFRVIDGGMALHQPPCKENETEREAGKSIYKIVKEK 413

Db 361 EEFNHWGGEDVEFGYRLFRKGCFFRVIDGGMALHQPPCKENETEREAGKSIYKIVKEK 420

QY 414 VPYIYKLLPIEDSHTRIPVLSIYIPAYNCANYIQRVDSALNQVVDLEVCINDGST 473

Db 421 VPYIYKLLPIEDSHTRIPVLSIYIPAYNCANYIQRVDSALNQVVDLEVCINDGST 480

QY 474 DNTLEVINKLYGNPRVRIMSKPENGGIASNAVAASFAGYIYIGLDDDDYLPDPDAVELC 533

Db 481 DNTLEVINKLYGNPRVRIMSKPENGGIASNAVAASFAGYIYIGLDDDDYLPDPDAVELC 540

QY 534 LKEFLKDKTLACVYTTNNRVNPDGSLIANGYNKPEFSREKLTWTAMTAHHEFMETIRAWHL 593

Db 541 LKEFLKDKTLACVYTTNNRVNPDGSLIANGYNKPEFSREKLTWTAMTAHHEFMETIRAWHL 600

QY 594 TDGFNENIENAVDYMFLKLVSEVGFKHLNKCYNRVLHGDNTSTIKKLGIOKKNHFVVVN 653
DB 601 TDGFNENIENAVDYMFLKLVSEVGFKHLNKCYNRVLHGDNTSTIKKLGIOKKNHFVVVN 660
QY 654 QSLNRQGINNYNDKFDOLDDESRYIFNKTAEYQEMDMLDKLKLIONKDAKIAVSIYFYP 713
DB 661 QSLNRQGINNYNDFDOLDDESRYIFNKTAEYQBEIDILDKIKIIONKDAKIAVSIYFYP 720
QY 714 NTLNGLVKKLNNIEYKNKIFVILHVDKNHLPDPIKEILAFYKHQVNNLLNNDISYY 773
DB 721 NTLNGLVKKLNNIEYKNKIFVILHVDKNHLPDPIKEILAFYKHQVNNLLNNDISYY 780
QY 774 TSNRLIKTEAHLNKLNSQLNCEYIIFDNHDSLFVKNDSYAYMKYDVGMMFSALTH 833
DB 781 TSNRLIKTEAHLNKLNSQLNCEYIIFDNHDSLFVKNDSYAYMKYDVGMMFSALTH 840
QY 834 DWIEKINAHPPFKLIKTYFNDNLSRMNVKASQGMFKYALPHELITTIKEVITSQOS 893
DB 841 DWIEKINAHPPFKLIKTYFNDNLSRMNVKASQGMFKYALPHELITTIKEVITSQOS 900
QY 894 IDSVPYNTEDIWQFALLILEKKTGHVFNKTSTLTYPWPKRLQWTHNEQIOSAKKGENI 953
DB 901 IDSVPYNTEDIWQFALLILEKKTGHVFNKTSTLTYPWPKRLQWTHNEQIOSAKKGENI 960
QY 954 PVNKFIIINSITL 965
DB 961 PVNKFIIINSITL 972

RESULT 13
US-09-791-537-56855
; Sequence 56855, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 56855
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-09-791-537-56855

Query Match 88.3%; Score 4493.5; DB 21; Length 972;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 844; Conservative 61; Mismatches 60; Indels 7; Gaps 2;

QY 1 MNTLSQAIKAYNSNDYELALKLFESAETYGRIKIVEFOLIKCKEKL----STNS---YVS 53
DB 1 MNTLSQAIKAYNSNDYELALKLFESAETYGRIKIVEFOLIKCKEKLKSAHPVSNSAHLNVN 60
QY 54 EDKNSVCDSSLDIATQALLSNVKKLTLSSEKSNLKNKWSITGKKSNAEIRKVELVP 113
DB 61 KEKVNVCDSPLDIATQALLSNVKKLVLSDEKNTLKNKKWLTEKKSENAEVRVALVP 120
QY 114 KDFPKDLVLAFLPDHVNDFTYWKNRKSGLGTPVKNKNTGLSIIITFNRSRILDTLACL 173
DB 121 KDFPKDLVLAFLPDHVNDFTYWKNRKSGLGTPVKNKNTGLSIIITFNRSRILDTLACL 180
QY 174 VNQKTYPFVYVAVDGSKNLTTVQKYEOKLDIKYVROKDYQOLCAVRNLGLRTAKY 233
DB 181 VNQKTYPFVYVAVDGSKNLTTVQKYEOKLDIKYVROKDYQOLCAVRNLGLRTAKY 240
QY 234 DVFSLDCDMPAQQLWVHSYLTLELDDNDIVLIGPRKYVDVTHNITAEQFLNDPYLIESLP 293
DB 241 DFIGLDCDMPAQQLWVHSYLTLELDDNDIVLIGPRKYVDVTHNITAEQFLNDPYLIESLP 300

QY 294 ETATNNPSITSKGNISLDWRLEHFKKTDNLRCLDSFERYFVAGNVAFSEKWLKNYGVWD 353
DB 301 EVKTNNSVAAKGEVTVSLDWRLEQFEKTEENLRUSDSPFRFAAGNVAFRAKWLKNSGFFD 360
QY 354 EEFNHWGGEDVEFGYRLFAGCGFFRVIDGGMATIHQBPFGKENEETEREAGKSITPLKIVKEK 413
DB 361 EEFNHWGGEDVEFGYRLFAGCGFFRVIDGGMATIHQBPFGKENEETEREAGKSITPLKIVKEK 420
QY 414 VPIYTRKLLPIEDSHIRPLVSIYIPAYNCANYIQCVDSSALNQTVVDLEVCINDGST 473
DB 421 VPIYTRKLLPIEDSHIRPLVSIYIPAYNCANYIQCVDSSALNQTVVDLEVCINDGST 480
QY 474 DNTLEVINLKYGNPNRVRITMSKPNGGIASASNAVAFKGYITGOLDSDDDYLPDPAVELC 533
DB 481 DNTLEVINLKYGNPNRVRITMSKPNGGIASASNAVAFKGYITGOLDSDDDYLPDPAVELC 540
QY 534 LKEFLKDKTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHHPMFTIRAWHL 593
DB 541 LKEFLKDKTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHHPMFTIRAWHL 600
QY 594 TDGFNENIENAVDYMFLKLVSEVGFKHLNKCYNRVLHGDNTSTIKKLGIOKKNHFVVVN 653
DB 601 TDGFNENIENAVDYMFLKLVSEVGFKHLNKCYNRVLHGDNTSTIKKLGIOKKNHFVVVN 660
QY 654 QSLNRQGINNYNDKFDOLDDESRYIFNKTAEYQBEIDILDKIKIIONKDAKIAVSIYFYP 713
DB 661 QSLNRQGINNYNDFDOLDDESRYIFNKTAEYQBEIDILDKIKIIONKDAKIAVSIYFYP 720
QY 714 NTLNGLVKKLNNIEYKNKIFVILHVDKNHLPDPIKEILAFYKHQVNNLLNNDISYY 773
DB 721 NTLNGLVKKLNNIEYKNKIFVILHVDKNHLPDPIKEILAFYKHQVNNLLNNDISYY 780
QY 774 TSNRLIKTEAHLNKLNSQLNCEYIIFDNHDSLFVKNDSYAYMKYDVGMMFSALTH 833
DB 781 TSNRLIKTEAHLNKLNSQLNCEYIIFDNHDSLFVKNDSYAYMKYDVGMMFSALTH 840
QY 834 DWIEKINAHPPFKLIKTYFNDNLSRMNVKASQGMFKYALPHELITTIKEVITSQOS 893
DB 841 DWIEKINAHPPFKLIKTYFNDNLSRMNVKASQGMFKYALPHELITTIKEVITSQOS 900
QY 894 IDSVPYNTEDIWQFALLILEKKTGHVFNKTSTLTYPWPKRLQWTHNEQIOSAKKGENI 953
DB 901 IDSVPYNTEDIWQFALLILEKKTGHVFNKTSTLTYPWPKRLQWTHNEQIOSAKKGENI 960
QY 954 PVNKFIIINSITL 965
DB 961 PVNKFIIINSITL 972

RESULT 14
US-10-011-768B-9
; Sequence 9, Application US/10011768B
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H.
; APPLICANT: DeAngelis, Paul
; APPLICANT: Kumari, Kshama
; TITLE OF INVENTION: Hyaluronan Synthase Gene and Uses Thereof
; FILE REFERENCE: 3554.011
; CURRENT APPLICATION NUMBER: US/10/011,768B
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 09/178,851
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: US 60/064,435
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Pasteurella Multocida
US-10-011-768B-9

Query Match		88.2%;	Score 4490.5;	DB 24;	Length 972;
Best Local Similarity		86.7%;	Pred. No. 0;		
Matches 843;		Conservative 62;	Mismatches 60;	Indels 7;	Gaps 2;
y	1	MNTLSQAIKAYNSNDYELALKLFESAEYGRKIVFEQIIKCKEKL-----STNS---YVS 53			
b	1	MNTLSQAIKAYNSNDYQALALKLFESAEYGRKIVFEQITCKEKLSAHPSVNSAHLNVN 60			
y	54	EDKNSVCDSDIATQALLSNVKKLTSESEKNSLKNKWSITGKSENAEIRKVELVP 113			
b	61	KEEKVNVCDSPDIATQALLSNVKKLVSDSEKNTLKNKKLLTEKKSENAEVRVALVP 120			
y	114	KDPKDLVAPLPHVNDFTWYKRRKSLGKIPVKNKIGLSIIPTFNRSRIIDITLACL 173			
b	121	KDPKDLVAPLPHVNDFTWYKRRKRLGKIPKHQHVGLSIIVTTFNRPAILSLITLACL 180			
y	174	VNQKNTYPFVAVDGSKENLTIYQKYEQKLDIKYVRQKDYQOLCAVRNGLRTAKY 233			
b	181	VNQKTHYFVIVTDGSGEDLSPIIRQYENKLDIRYVRQKNGFQASARNMGLRLAKY 240			
y	234	DFVSLDCCDAPQOLWVHSYLTLELDDNDIVLIGPRKYVDTHNITAEQFLNDPVLIESLP 293			
b	241	DFIGLLDCDAPNPLWVHSYVAELLEDLTIIGPRXIYDTHQIDPKDFLNNASLLES 300			
y	294	ETATNNPISITSGNISLDRLEHFKKTDNLRLCDSPPRYFVAGNVAFSEKWLNVKVGFD 353			
b	301	EVKTNNVAAKGEVTVSLDWRLQEFKTEENLRSLSDSPFRFAAGNVAFSAKWLKNSGFFD 360			
y	354	EEFNHWGGEDEVFYRLFAKCGCFRVIDGMAIHQEPGPKENETERAGKSITIKYVEK 413			
b	361	EEFNHWGGEDEVFYRLFRYGSFFKTIDGIMAYHQEPGPKENETDREAGKNITLDMREK 420			
y	414	VPYIRKLLPTEDSHRIPLVSIYIPAYNCANYIQRVDSALNQTVVDLEVCINCDSGT 473			
b	421	VPYIRKLLPTEDSHINRPLVSIYIPAYNCANYIQRVDSALNQTVVDLEVCINCDSGT 480			
y	474	DNTLEVINKLYGNPRVIRMSKPNGGIASASNAAVSFAGYIYIGOLDSDDYLEPDAVELC 533			
b	481	DNTLEVINKLYGNPRVIRMSKPNGGIASASNAAVSFAGYIYIGOLDSDDYLEPDAVELC 540			
y	534	LKEFLKDKTLACVYTTNRNVPDGLSIANGYNWPFESREKLTMTAMIAHFRMTTIRAWHL 593			
b	541	LKEFLKDKTLACVYTTNRNVPDGLSIANGYNWPFESREKLTMTAMIAHFRMTTIRAWHL 600			
y	594	TGDFENIENAVDYDMFLKSEVGKFKHLNKCICYNRVLHGDNTSIKLGIOKKNHFVVVN 653			
b	601	TGDFNEKIENAVDYDMFLKSEVGKFKHLNKCICYNRVLHGDNTSIKLGIOKKNHFVVVN 660			
y	654	QSLNRQGINYNYDKFDDLESRYIFNKTAEQEEMDKLKLQNKDAKIAVSIYFP 713			
b	661	QSLNRQGITTYNYDFDDESRYIFNKTAEQEEIDILKILQNKDAKIAVSIYFP 720			
y	714	NFLNGLVKLNNIIEYNKNIFVILHVDKNHLLTPDKKEILAFYHKHQNILLNNDISY 773			
b	721	NFLNGLVKLNNIIEYNKNIFVILHVDKNHLLTPDKKEILAFYHKHQNILLNNDISY 780			
y	774	TSNRLIKTEAHLNKLNSQLNCEYIIFDNHDSLFVKNDSYAYMKYDVGMNFSALTH 833			
b	781	TSNRLIKTEAHLNKLNSQLNCEYIIFDNHDSLFVKNDSYAYMKYDVGMNFSALTH 840			
y	834	DWIEKINAPPPKLIKITYFNDNDRSMNVKASQGMFKYALPHELLTIIKEVITSCQS 893			
b	841	DWIEKINAPPPKLIKITYFNDNDRSMNVKASQGMFKYALPHELLTIIKEVITSCQS 900			
y	894	IDSVPEYNTEDLWFQFALLILEKKTGHVFNKTSTLTYPWPERKLTQNTNEQISAKKGENI 953			
b	901	IDSVPEYNTEDLWFQFALLILEKKTGHVFNKTSTLTYPWPERKLTQNTNEQISAKKGENI 960			
y	954	PVNKFIINSITL 965			
b	961	PVNKFIINSITL 972			

RESULT 15
US-10-011-771B-9
; Sequence 9, Application US/10011771B
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H.
; APPLICANT: DeAngelis, Paul
; APPLICANT: Kumari, Kshama
; TITLE OF INVENTION: Hyaluronan Synthase Gene and Uses Thereof
; FILE REFERENCE: 3554.011
; CURRENT APPLICATION NUMBER: US/10/011.771B
; CURRENT FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: US 09/178,851
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: US 60/064,435
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Pastuerella Multocida
US-10-011-771B-9

Query Match	88.2%;	Score 4490.5;	DB 24;	Length 972;
Best Local Similarity	86.7%;	Pred. No. 0;		
Matches 843;	Conservative 62;	Mismatches 60;	Indels 7;	Gaps 2;
Qy	1	MNTLSQAIKAYNSNDYELALKLFESAEYGRKIVFEQIIKCKEKL-----STNS---YVS 53		
Db	1	MNTLSQAIKAYNSNDYQALALKLFESAEYGRKIVFEQITCKEKLSAHPSVNSAHLNVN 60		
Qy	54	EDKNSVCDSDIATQALLSNVKKLTISESEKNSLKNKWSITGKSENAEIRKVELVP 113		
Db	61	KEEKVNVCDSPDIATQALLSNVKKLVSDSEKNTLKNKKLLTEKKSENAEVRVALVP 120		
Qy	114	KDPKDLVAPLPHVNDFTWYKRRKSLGKIPVKNKIGLSIIPTFNRSRIIDITLACL 173		
Db	121	KDPKDLVAPLPHVNDFTWYKRRKRLGKIPKHQHVGLSIIVTTFNRPAILSLITLACL 180		
Qy	174	VNQKNTYPFVAVDGSKENLTIYQKYEQKLDIKYVRQKDYQOLCAVRNGLRTAKY 233		
Db	181	VNQKTHYFVIVTDGSGEDLSPIIRQYENKLDIRYVRQKNGFQASARNMGLRLAKY 240		
Qy	234	DFVSLDCCDAPQOLWVHSYLTLELDDNDIVLIGPRKYVDTHNITAEQFLNDPVLIESLP 293		
Db	241	DFIGLLDCDAPNPLWVHSYVAELLEDLTIIGPRXIYDTHQIDPKDFLNNASLLES 300		
Qy	294	ETATNNPISITSGNISLDRLEHFKKTDNLRLCDSPPRYFVAGNVAFSEKWLNVKVGFD 353		
Db	301	EVKTNNVAAKGEVTVSLDWRLQEFKTEENLRSLSDSPFRFAAGNVAFSAKWLKNSGFFD 360		
Qy	354	EEFNHWGGEDEVFYRLFAKCGCFRVIDGMAIHQEPGPKENETERAGKSITIKYVEK 413		
Db	361	EEFNHWGGEDEVFYRLFRYGSFFKTIDGIMAYHQEPGPKENETDREAGKNITLDMREK 420		
Qy	414	VPYIRKLLPTEDSHRIPLVSIYIPAYNCANYIQRVDSALNQTVVDLEVCINCDSGT 473		
Db	421	VPYIRKLLPTEDSHINRPLVSIYIPAYNCANYIQRVDSALNQTVVDLEVCINCDSGT 480		
Qy	474	DNTLEVINKLYGNPRVIRMSKPNGGIASASNAAVSFAGYIYIGOLDSDDYLEPDAVELC 533		
Db	481	DNTLEVINKLYGNPRVIRMSKPNGGIASASNAAVSFAGYIYIGOLDSDDYLEPDAVELC 540		
Qy	534	LKEFLKDKTLACVYTTNRNVPDGLSIANGYNWPFESREKLTMTAMIAHFRMTTIRAWHL 593		
Db	541	LKEFLKDKTLACVYTTNRNVPDGLSIANGYNWPFESREKLTMTAMIAHFRMTTIRAWHL 600		
Qy	594	TGDFENIENAVDYDMFLKSEVGKFKHLNKCICYNRVLHGDNTSIKLGIOKKNHFVVVN 653		
Db	601	TGDFNEKIENAVDYDMFLKSEVGKFKHLNKCICYNRVLHGDNTSIKLGIOKKNHFVVVN 660		
Qy	654	QSLNRQGINYNYDKFDDDESRYIFNKTAEYQEMDMKDLKLIQNKDAKIAVSIFYP 713		

Db	661	QSLNRQ	GITYNYDEFDLDESRYIFNKTAEQEIDILDKDIQNKDAKIAVSIFYP	720
Qy	714	NTLGLVKKLN	IIEYNKNIFVILHVDKNHLPDIKKKEILAFYKHQVNIILNNDISY	773
Db	721	NTLGLVKKLN	IIEYNKNIFVILHVDKNHLPDIKKKEILAFYKHQVNIILNNDISY	780
Qy	774	TSNRLIKTEA	HLNKLNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMNFSALTH	833
Db	781	TSNRLIKTEA	HLNKLNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMNFSALTH	840
Qy	834	DWIEKINAH	PPFKLIKTYFNDNDRSMNVKGASQGMFKYALPHELLTIIKEVITSCQS	893
Db	841	DWIEKINAH	PPFKLIKTYFNDNDRSMNVKGASQGMFKYALPHELLTIIKEVITSCQS	900
Qy	894	IDSYPEYNT	EDIWQFALLILEKKTGHVFNKTSTLTYPWPKLQWNEQIOSAKKGENI	953
Db	901	IDSYPEYNT	EDIWQFALLILEKKTGHVFNKTSTLTYPWPKLQWNEQIOSAKKGENI	960
Qy	954	PVKNFIINS	ITL 965	
Db	961	PVKNFIINS	ITL 972	

Search completed: January 4, 2003, 02:36:14
Job time : 405 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2003, 02:24:29 ; Search time 562 Seconds
(without alignments)
121.219 Million cell updates/sec

Title: US-09-842-484A-2
Perfect score: 5089
Sequence: 1 MNTLSQAIKAYNSNDYELAL.....SAKKGNIPIVKNKFIINSITL 965

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 258925 seqs, 70596210 residues
Total number of hits satisfying chosen parameters: 258925

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5089	100.0	965	5	US-09-842-484A-2
2	5067	99.6	965	5	US-09-842-484A-4
3	5067	99.6	965	6	US-10-217-613-3
4	5067	99.6	965	6	US-10-217-613-8
5	4604.5	90.5	972	6	US-10-217-613-9
6	4490.5	88.2	972	5	US-09-469-200D-9
7	4486.5	88.2	972	6	US-10-309-560-8
8	3102.5	61.0	703	6	US-10-217-613-1
9	3102.5	61.0	703	6	US-10-217-613-7
10	2256.5	44.3	686	6	US-10-216-289-2
11	2256.5	44.3	686	6	US-10-216-289-4
12	333.5	6.6	842	5	US-09-134-000C-5749
13	299	5.9	331	6	US-10-264-213-152
14	262.5	5.2	534	5	US-09-134-000C-5087
15	248.5	4.9	330	5	US-09-134-000C-6392
16	240.5	4.7	1056	5	US-09-134-000C-5086
17	227	4.5	727	6	US-10-092-411A-4067
18	210	4.1	348	6	US-10-096-129-3
19	210	4.1	348	6	US-10-096-129-8
20	209	4.1	337	6	US-10-096-129-5
21	208.5	4.1	270	6	US-10-303-161-39
22	208.5	4.1	270	6	US-10-303-162-39
23	208.5	4.1	270	6	US-10-303-118-39
24	208.5	4.1	270	6	US-10-303-128-39
25	208.5	4.1	270	6	US-10-303-134-39
26	207	4.1	301	6	US-10-303-161-27

27	207	4.1	301	6	US-10-303-162-27	Sequence 27, Appl
28	207	4.1	301	6	US-10-303-118-27	Sequence 27, Appl
29	207	4.1	301	6	US-10-303-128-27	Sequence 27, Appl
30	207	4.1	301	6	US-10-303-134-27	Sequence 27, Appl
31	205	4.0	389	6	US-10-303-161-34	Sequence 34, Appl
32	205	4.0	389	6	US-10-303-162-34	Sequence 34, Appl
33	205	4.0	389	6	US-10-303-118-34	Sequence 34, Appl
34	205	4.0	389	6	US-10-303-128-34	Sequence 34, Appl
35	205	4.0	389	6	US-10-303-134-34	Sequence 34, Appl
36	204	4.0	651	6	US-10-142-143-6	Sequence 6, Appl
37	200.5	3.9	303	6	US-10-303-161-29	Sequence 29, Appl
38	200.5	3.9	303	6	US-10-303-162-29	Sequence 29, Appl
39	200.5	3.9	303	6	US-10-303-118-29	Sequence 29, Appl
40	200.5	3.9	303	6	US-10-303-128-29	Sequence 29, Appl
41	200.5	3.9	303	6	US-10-303-134-29	Sequence 29, Appl
42	200	3.9	617	6	US-10-142-143-2	Sequence 2, Appl
43	199	3.9	297	6	US-10-303-161-31	Sequence 31, Appl
44	199	3.9	297	6	US-10-303-162-31	Sequence 31, Appl
45	199	3.9	297	6	US-10-303-118-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1
US-09-842-484A-2
; Sequence 2, Application US/09842484A
; GENERAL INFORMATION:
; APPLICANT: DEANGELIS, PAUL L.
; TITLE OF INVENTION: CHONDROITIN SYNTHASE GENE AND METHODS OF MAKING AND USING SAM
; FILE REFERENCE: 4605.003
; CURRENT APPLICATION NUMBER: US/09/842,484A
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 965
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-09-842-484A-2

Query Match	100.0%;	Score	5089;	DB	5;	Length	965;							
Best Local Similarity	100.0%;	Pred. No.	0;											
Matches	965;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;					
QY	1	MNTLSQA	IKAYNSNDYELAL	KLF	PKSAET	YGRKIV	EQIIK	KEKELSTNSV	SEDKKNSV	60				
Db	1	MNTLSQA	IKAYNSNDYELAL	KLF	PKSAET	YGRKIV	EQIIK	KEKELSTNSV	SEDKKNSV	60				
QY	61	CDSSLDI	ATQLLLSNVK	KL	TLSESEKNSL	NKWSITG	KKSENAEIR	KVELVP	KDFPKDL	120				
Db	61	CDSSLDI	ATQLLLSNVK	KL	TLSESEKNSL	NKWSITG	KKSENAEIR	KVELVP	KDFPKDL	120				
QY	121	VLAPLPH	VNDFTWYK	NRKKS	LIGIK	PKVKN	IGLSII	PTFNRS	RILDTIT	LACLVNQ	TNY	180		
Db	121	VLAPLPH	VNDFTWYK	NRKKS	LIGIK	PKVKN	IGLSII	PTFNRS	RILDTIT	LACLVNQ	TNY	180		
QY	181	PFVYVAD	GSGKENLL	TI	VQKYEQ	KLDIKYVR	QDKYGYQLCA	VNRNLG	LRTAKY	DFVSILD	240			
Db	181	PFVYVAD	GSGKENLL	TI	VQKYEQ	KLDIKYVR	QDKYGYQLCA	VNRNLG	LRTAKY	DFVSILD	240			
QY	241	CDMAPOOL	WVHSYLTLEL	ENDIV	LIGPRKY	VDTHNT	ATAEQFLNDP	YLIESLP	TATNN	300				
Db	241	CDMAPOOL	WVHSYLTLEL	ENDIV	LIGPRKY	VDTHNT	ATAEQFLNDP	YLIESLP	TATNN	300				
QY	301	PSITSGK	NTSLDWRL	EHFK	KTDN	LR	LCDS	SPRYFVAGN	VAFSK	EWLKNK	GVWFDE	EENHWG	360	
Db	301	PSITSGK	NTSLDWRL	EHFK	KTDN	LR	LCDS	SPRYFVAGN	VAFSK	EWLKNK	GVWFDE	EENHWG	360	
QY	361	GEDVFEY	GRILFAK	GCFFR	YID	GGMA	IHQEP	PGKEN	ETRE	AGKSIT	LKIV	KEKVP	YIYRK	420

```
Db 361 GEDVEGYRLFAKCGFRVLDGGMALHQBPPGKNETEREAGKSITLKIIVKEKVPYIYRK 420
QY 421 LLPEDSHIHRIPLVSIYIPAYNCANYIQCVDVSALNQTVVDLEVCICNDGSDNTLEVI 480
Db 421 LLPEDSHIHRIPLVSIYIPAYNCANYIQCVDVSALNQTVVDLEVCICNDGSDNTLEVI 480
QY 481 NKLYGNPRVRIMSKPNGGKTASASNAVSAFAGYITGOLSDDDYLEDPAVELCLKEFLKD 540
Db 481 NKLYGNPRVRIMSKPNGGKTASASNAVSAFAGYITGOLSDDDYLEDPAVELCLKEFLKD 540
QY 541 KTLACVYTTNRNVPDGLSIANGYNWPEFSREKLTITAMIAHHFRMFTIRAWHLTDGFEN 600
Db 541 KTLACVYTTNRNVPDGLSIANGYNWPEFSREKLTITAMIAHHFRMFTIRAWHLTDGFEN 600
QY 601 IENAVDYDMFLKSEVGKFKHLNKCYNRVLHGDNTSIKKLGIOQKKNHVVVQSLNRQG 660
Db 601 IENAVDYDMFLKSEVGKFKHLNKCYNRVLHGDNTSIKKLGIOQKKNHVVVQSLNRQG 660
QY 661 INYNYDKFDLDESRYIFNKTAEYOEDMDLKDQLIONKDAKIAVSIFYPNTLNGLV 720
Db 661 INYNYDKFDLDESRYIFNKTAEYOEDMDLKDQLIONKDAKIAVSIFYPNTLNGLV 720
QY 721 KKLNNIIEYKNIIFVILHVDKNHLPDIKKEILAFYHKHQNVLNNDISYYSNRLIK 780
Db 721 KKLNNIIEYKNIIFVILHVDKNHLPDIKKEILAFYHKHQNVLNNDISYYSNRLIK 780
QY 781 TEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDYGMMFSALTDHWDIEKIN 840
Db 781 TEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDYGMMFSALTDHWDIEKIN 840
QY 841 AHPPEFKLITKYFNDNDRSMNVKASOGMPKMYALPHELLTIIKEVITSCQSIDSVPEY 900
Db 841 AHPPEFKLITKYFNDNDRSMNVKASOGMPKMYALPHELLTIIKEVITSCQSIDSVPEY 900
QY 901 NTEDIWFQFALLILEKKTGHVFNKSTLTLYMPWERKLOWTNEQIQSAKKGENIPVNFII 960
Db 901 NTEDIWFQFALLILEKKTGHVFNKSTLTLYMPWERKLOWTNEQIQSAKKGENIPVNFII 960
QY 961 NSITL 965
Db 961 NSITL 965
```

```
RESULT 2
US-09-842-484A-4
; Sequence 4, Application US/09842484A
; GENERAL INFORMATION:
; APPLICANT: DEANGELIS, PAUL L.
; TITLE OF INVENTION: CHONDROITIN SYNTHASE GENE AND METHODS OF MAKING AND USING SAME
; FILE REFERENCE: 4605.003
; CURRENT APPLICATION NUMBER: US/09/842.484A
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ. ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ. ID NO 4
; LENGTH: 965
; TYPE: PRF
; ORGANISM: Pasteurella multocida
US-09-842-484A-4

Query Match 99.6%; Score 5067; DB 5; Length 965;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 960; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MNTLSQAIKAYNSDYELAKLIFKSAETYGRIKVEFQIIKCKELSTNSYVSEDKKNSV 60
Db 1 MNTLSQAIKAYNSDYELAKLIFKSAETYGRIKVEFQIIKCKELSTNSYVSEDKKNSV 60
QY 61 CDSSLDIATQLLSNVKLTLSSEKNSLKNKWSITGKKSNAERKVELYVPKDPKDL 120
Db 61 CDSSLDIATQLLSNVKLTLSSEKNSLKNKWSITGKKSNAERKVELYVPKDPKDL 120
```

```
QY 121 VLAPLDPHVNDFTWYKNRKSLSGIKPVNKNIGLSIIIIPTFNRSRIILDITLACLVNQKTN 180
Db 121 VLAPLDPHVNDFTWYKNRKSLSGIKPVNKNIGLSIIIIPTFNRSRIILDITLACLVNQKTN 180
QY 181 PFEVVVADDSKKNLLTIVQYEQKLDIKYVRQKDYQYLCVARNLGLRTAKYDFVSTLD 240
Db 181 PFEVVVADDSKKNLLTIVQYEQKLDIKYVRQKDYQYLCVARNLGLRTAKYDFVSTLD 240
QY 241 CDMAPOQLWVHSYLTLELNDIVLIGPRKYVDTHNTATAEQFLNDPYLESLPETATN 300
Db 241 CDMAPOQLWVHSYLTLELNDIVLIGPRKYVDTHNTATAEQFLNDPYLESLPETATN 300
QY 301 PSITSKGNISLDWRLEHFKTKTDLNRLCDSPPRYFVAGNVAFSKEWLNKVGWDFEENHWG 360
Db 301 PSITSKGNISLDWRLEHFKTKTDLNRLCDSPPRYFVAGNVAFSKEWLNKVGWDFEENHWG 360
QY 361 GEDVEGYRLFAKCGFRVLDGGMALHQBPPGKNETEREAGKSITLKIIVKEKVPYIYRK 420
Db 361 GEDVEGYRLFAKCGFRVLDGGMALHQBPPGKNETEREAGKSITLKIIVKEKVPYIYRK 420
QY 421 LLPEDSHIHRIPLVSIYIPAYNCANYIQCVDVSALNQTVVDLEVCICNDGSDNTLEVI 480
Db 421 LLPEDSHIHRIPLVSIYIPAYNCANYIQCVDVSALNQTVVDLEVCICNDGSDNTLEVI 480
QY 481 NKLYGNPRVRIMSKPNGGKTASASNAVSAFAGYITGOLSDDDYLEDPAVELCLKEFLKD 540
Db 481 NKLYGNPRVRIMSKPNGGKTASASNAVSAFAGYITGOLSDDDYLEDPAVELCLKEFLKD 540
QY 541 KTLACVYTTNRNVPDGLSIANGYNWPEFSREKLTITAMIAHHFRMFTIRAWHLTDGFEN 600
Db 541 KTLACVYTTNRNVPDGLSIANGYNWPEFSREKLTITAMIAHHFRMFTIRAWHLTDGFEN 600
QY 601 IENAVDYDMFLKSEVGKFKHLNKCYNRVLHGDNTSIKKLGIOQKKNHVVVQSLNRQG 660
Db 601 IENAVDYDMFLKSEVGKFKHLNKCYNRVLHGDNTSIKKLGIOQKKNHVVVQSLNRQG 660
QY 661 INYNYDKFDLDESRYIFNKTAEYOEDMDLKDQLIONKDAKIAVSIFYPNTLNGLV 720
Db 661 INYNYDKFDLDESRYIFNKTAEYOEDMDLKDQLIONKDAKIAVSIFYPNTLNGLV 720
QY 721 KKLNNIIEYKNIIFVILHVDKNHLPDIKKEILAFYHKHQNVLNNDISYYSNRLIK 780
Db 721 KKLNNIIEYKNIIFVILHVDKNHLPDIKKEILAFYHKHQNVLNNDISYYSNRLIK 780
QY 781 TEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDYGMMFSALTDHWDIEKIN 840
Db 781 TEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDYGMMFSALTDHWDIEKIN 840
QY 841 AHPPEFKLITKYFNDNDRSMNVKASOGMPKMYALPHELLTIIKEVITSCQSIDSVPEY 900
Db 841 AHPPEFKLITKYFNDNDRSMNVKASOGMPKMYALPHELLTIIKEVITSCQSIDSVPEY 900
QY 901 NTEDIWFQFALLILEKKTGHVFNKSTLTLYMPWERKLOWTNEQIQSAKKGENIPVNFII 960
Db 901 NTEDIWFQFALLILEKKTGHVFNKSTLTLYMPWERKLOWTNEQIQSAKKGENIPVNFII 960
QY 961 NSITL 965
Db 961 NSITL 965
```

```
RESULT 3
US-10-217-613-3
; Sequence 3, Application US/10217613
; GENERAL INFORMATION:
; APPLICANT: DEANGELIS, PAUL
; TITLE OF INVENTION: DNA ENCODING HYALURONAN SYNTHASE FROM PASTURELLA MULTOCIDA ANI
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 35341.081
; CURRENT APPLICATION NUMBER: US/10/217,613
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/283,402
```

```
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 965
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-217-613-3

Query Match          99.6%; Score 5067; DB 6; Length 965;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 960; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGRIKIVEFQIIKCKEKLSTNSYVSEDKKNSV 60
Db 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGRIKIVEFQIIKCKEKLSTNSYVSEDKKNSV 60

Qy 61 CDSSLDIATQLLSNVKLTISESEKNSLKNKWSITGKSENAEIRKVELVPKDFPKDL 120
Db 61 CDSSLDIATQLLSNVKLTISESEKNSLKNKWSITGKSENAEIRKVELVPKDFPKDL 120

Qy 121 VLAPLDHVNDFWYKNNRKSGLGKIPVKNKIGLSIIPTFNRSRIIDITLACLNVQKNTY 180
Db 121 VLAPLDHVNDFWYKNNRKSGLGKIPVKNKIGLSIIPTFNRSRIIDITLACLNVQKNTY 180

Qy 181 PFEVVVADGSKENLLTIVQKYEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKYDFVSILD 240
Db 181 PFEVVVADGSKENLLTIVQKYEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKYDFVSILD 240

Qy 241 CDMAPOOLWVHSLYTELLEDNDIVLIGPRKYVDTHNITAEQFLNDPYLESPETATNNN 300
Db 241 CDMAPOOLWVHSLYTELLEDNDIVLIGPRKYVDTHNITAEQFLNDPYLESPETATNNN 300

Qy 301 PSITSGKNSLDWRLEHFKKTDNLRCLDSPFRYFVAGNVAFSKEWLNKVGWDFDEPNHWG 360
Db 301 PSITSGKNSLDWRLEHFKKTDNLRCLDSPFRYFVAGNVAFSKEWLNKVGWDFDEPNHWG 360

Qy 361 GEDVEFGYRLFAKGCFFRVIDGGMAYHQEPGKENETDREAGKSITLKIVKEKVPYIYRK 420
Db 361 GEDVEFGYRLFAKGCFFRVIDGGMAYHQEPGKENETDREAGKSITLKIVKEKVPYIYRK 420

Qy 421 LLPIEDSHIHRPLVSIYIPAYNCANYIQRCVDSALNQTVDVLEVCICNDGSTDNTLEVI 480
Db 421 LLPIEDSHIHRPLVSIYIPAYNCANYIQRCVDSALNQTVDVLEVCICNDGSTDNTLEVI 480

Qy 481 NKLYGNPNRVRIMSKPNGGIASASNAVSFAKGYITGQLDSDDYLEPDADAVELCLKEFLKD 540
Db 481 NKLYGNPNRVRIMSKPNGGIASASNAVSFAKGYITGQLDSDDYLEPDADAVELCLKEFLKD 540

Qy 541 KTLACVYTTNRNVNPDGSLIANGYNWPEPSREKLTAMTAHFRMPTIRAWHLTDGFNEN 600
Db 541 KTLACVYTTNRNVNPDGSLIANGYNWPEPSREKLTAMTAHFRMPTIRAWHLTDGFNEN 600

Qy 601 IENAVDYDMFLKSLSEVGKFKLNKICYNRVLHGDNTSIKKLGLOKKNHFVVVQNSLNROG 660
Db 601 IENAVDYDMFLKSLSEVGKFKLNKICYNRVLHGDNTSIKKLGLOKKNHFVVVQNSLNROG 660

Qy 661 INYNYDKPDDIDESRKYIFNKTAEQEEMDLKDLKLQNKDQAKTAVSIFYPNTUNGLV 720
Db 661 INYNYDKPDDIDESRKYIFNKTAEQEEMDLKDLKLQNKDQAKTAVSIFYPNTUNGLV 720

Qy 721 KKLNNIIEYKNKIFVILHVDKNHLTPDIKKELAFYHKHQVNILLNNDISYVTSNRLIK 780
Db 721 KKLNNIIEYKNKIFVILHVDKNHLTPDIKKELAFYHKHQVNILLNNDISYVTSNRLIK 780

Qy 781 TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMNFSALTHDWTIEKIN 840
Db 781 TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMNFSALTHDWTIEKIN 840
```

```
Qy 841 AHPPFKKLIKTYFNDNDLRSMNVKGASQGMFKYALPHELLTIIKEVITSQSIDSVPEY 900
Db 841 AHPPFKKLIKTYFNDNDLRSMNVKGASQGMFKYALPHELLTIIKEVITSQSIDSVPEY 900

Qy 901 NTEDIWFOFALLILEKKTGHVFNKSTLTLYMPWERKLQWNTNEQIQSAKKGENIPVKNKFI 960
Db 901 NTEDIWFOFALLILEKKTGHVFNKSTLTLYMPWERKLQWNTNEQIQSAKKGENIPVKNKFI 960

Qy 961 NSITL 965
Db 961 NSITL 965

RESULT 4
US-10-217-613-8
; Sequence 8, Application US/10217613
; GENERAL INFORMATION:
; APPLICANT: D'ANGELIS, PAUL
; TITLE OF INVENTION: DNA ENCODING HYALURONAN SYNTHASE FROM PASTURELLA MULTOCIDA AT
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 35541.081
; CURRENT APPLICATION NUMBER: US/10/217,613
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/283,402
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 965
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-217-613-8

Query Match          99.6%; Score 5067; DB 6; Length 965;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 960; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGRIKIVEFQIIKCKEKLSTNSYVSEDKKNSV 60
Db 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGRIKIVEFQIIKCKEKLSTNSYVSEDKKNSV 60

Qy 61 CDSSLDIATQLLSNVKLTISESEKNSLKNKWSITGKSENAEIRKVELVPKDFPKDL 120
Db 61 CDSSLDIATQLLSNVKLTISESEKNSLKNKWSITGKSENAEIRKVELVPKDFPKDL 120

Qy 121 VLAPLDHVNDFWYKNNRKSGLGKIPVKNKIGLSIIPTFNRSRIIDITLACLNVQKNTY 180
Db 121 VLAPLDHVNDFWYKNNRKSGLGKIPVKNKIGLSIIPTFNRSRIIDITLACLNVQKNTY 180

Qy 181 PFEVVVADGSKENLLTIVQKYEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKYDFVSILD 240
Db 181 PFEVVVADGSKENLLTIVQKYEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKYDFVSILD 240

Qy 241 CDMAPOOLWVHSLYTELLEDNDIVLIGPRKYVDTHNITAEQFLNDPYLESPETATNNN 300
Db 241 CDMAPOOLWVHSLYTELLEDNDIVLIGPRKYVDTHNITAEQFLNDPYLESPETATNNN 300

Qy 301 PSITSGKNSLDWRLEHFKKTDNLRCLDSPFRYFVAGNVAFSKEWLNKVGWDFDEPNHWG 360
Db 301 PSITSGKNSLDWRLEHFKKTDNLRCLDSPFRYFVAGNVAFSKEWLNKVGWDFDEPNHWG 360

Qy 361 GEDVEFGYRLFAKGCFFRVIDGGMAYHQEPGKENETDREAGKSITLKIVKEKVPYIYRK 420
Db 361 GEDVEFGYRLFAKGCFFRVIDGGMAYHQEPGKENETDREAGKSITLKIVKEKVPYIYRK 420

Qy 421 LLPIEDSHIHRPLVSIYIPAYNCANYIQRCVDSALNQTVDVLEVCICNDGSTDNTLEVI 480
Db 421 LLPIEDSHIHRPLVSIYIPAYNCANYIQRCVDSALNQTVDVLEVCICNDGSTDNTLEVI 480
```

```
QY 481 NKLYGNPRVRIMSKPNGGIASASNAVSPAKGYIIGQLSDDDYLEPDAVELCLKEFLKD 540
Db 481 NKLYGNPRVRIMSKPNGGIASASNAVSPAKGYIIGQLSDDDYLEPDAVELCLKEFLKD 540
QY 541 KTLACVYTTNRNPNPDSGLIANGYNWPEFSREKLTAMIAHHPFMTIRAWHLTDGFEN 600
Db 541 KTLACVYTTNRNPNPDSGLIANGYNWPEFSREKLTAMIAHHPFMTIRAWHLTDGFEN 600
QY 601 IENAVDYDMFLKILSEVGKFKHLNKCYNRVLHGDNTSIKKLGIOKKNHFFVVNQSLNRQ 660
Db 601 IENAVDYDMFLKILSEVGKFKHLNKCYNRVLHGDNTSIKKLGIOKKNHFFVVNQSLNRQ 660
QY 661 INYNYDKFDLDESRYIENKTAEQEEMDLKDLKLNQKDAKIAVSIFYPNTLNGLV 720
Db 661 INYNYDKFDLDESRYIENKTAEQEEMDLKDLKLNQKDAKIAVSIFYPNTLNGLV 720
QY 721 KKLNNIIEYNKNTFVILHLVDKNHLPDIKKEILAFYKHQVNLNNDISYVTSNRLIK 780
Db 721 KKLNNIIEYNKNTFVILHLVDKNHLPDIKKEILAFYKHQVNLNNDISYVTSNRLIK 780
QY 781 TEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSAETHDWIEKIN 840
Db 781 TEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSAETHDWIEKIN 840
QY 841 AHPFPKLIKTYFNDNDRSNVKGASQGMFKYALPHELLTIIKEVITSCQSIDSVPEY 900
Db 841 AHPFPKLIKTYFNDNDRSNVKGASQGMFKYALPHELLTIIKEVITSCQSIDSVPEY 900
QY 901 NTEIDWFOFALLILEKKTGHVFNKTSITLYMPWERKLOWNEQIOSAKKGENIPVNFII 960
Db 901 NTEIDWFOFALLILEKKTGHVFNKTSITLYMPWERKLOWNEQIOSAKKGENIPVNFII 960
QY 961 NSITL 965
Db 961 NSITL 965

RESULT 5
US-10-217-613-9
; Sequence 9, Application US/10217613
; GENERAL INFORMATION:
; APPLICANT: DEANGELIS, PAUL
; TITLE OF INVENTION: DNA ENCODING HYALURONAN SYNTHASE FROM PASTURELLA MULTOCIDA AND US
; FILE REFERENCE: 35541.081
; CURRENT APPLICATION NUMBER: US/10/217,613
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/283,402
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (17)..(17)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (62)..(62)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (90)..(90)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (113)..(113)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (158)..(158)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (164)..(164)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (192)..(192)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (201)..(201)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (206)..(206)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (211)..(211)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (225)..(225)
; OTHER INFORMATION: either Phe or Tyr
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (233)..(233)
; OTHER INFORMATION: either Met or Leu
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (243)..(243)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (253)..(253)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (279)..(279)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (288)..(288)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (292)..(292)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (316)..(316)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (329)..(329)
; OTHER INFORMATION: either Phe or Tyr
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (340)..(340)
; OTHER INFORMATION: either Phe or Tyr
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (405)..(405)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
```

```
; NAME/KEY: MISC_FEATURE
; LOCATION: (439)..(439)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (744)..(744)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (952)..(952)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
US-10-217-613-9

Query Match      90.5%; Score 4604.5; DB 6; Length 972;
Best Local Similarity 89.9%; Pred. No. 0;
Matches 874; Conservative 29; Mismatches 62; Indels 7; Gaps 1;

Qy 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGKIVFQIICKKEKLSN-----YVS 53
Db 1 MNTLSQAIKAYNSNDYQALKLFKSAETYGKIVFQIICKKEKLSANPSVNEANLSVN 60
Qy 54 EDKNSVCDSSLDIATQALLSNVKKLTLSSEKNSLKNKWSITGKSENAEIRKVELVP 113
Db 61 EXEKNVCDSPDLIATQALLSNVKKLTLSXSEKNSLKNKWLITEKSENAEIRKVELVP 120
Qy 114 KDFPKDLVLAFLPDHVNDFTWYKNRKSLSGKPKVKNKIGLSIIPTFNRSLDITLACL 173
Db 121 KDFPKDLVLAFLPDHVNDFTWYKNRKSLSGKPKVKNKIGLSIIPTFNRSLDITLACL 180
Qy 174 VNQKTYPEVVDGSGKENLTTIVQYEQKLDITKYVRQKDYGYQLCAVRNLGLRTAKY 233
Db 181 VNQKTYPEVVDGSGQEXLLPLIXROYEXKLDITRYVRQKDYGYQACAAARXGLRLAKY 240
Qy 234 DFVSTLDCDMPAQOLWVHSYLTLELDDNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 293
Db 241 DFVSTLDCDMPAQOLWVHSYLTLELDDNDITIGPRKYVDTHNITAEQFLNDPYLIESLP 300
Qy 294 ETATNNPNSITSGNISLDRLEHFKTDNLRCLDPSERYFVAGNVAFSKWLNKVGWFD 353
Db 301 ETATNNPNSITSGNISLDRLEHFKTDNLRCLDPSERYFVAGNVAFSKWLNKVGWFD 360
Qy 354 EEFNHGGEDEVFGYRLFAKGCFFRVIDGMAIHQEPGKENETREAGKSTTLKIVREK 413
Db 361 EEFNHGGEDEVFGYRLFAKGCFFRVIDGMAIHQEPGKENETREAGKSTTLKIVREK 420
Qy 414 VPIYRKLLPIEDSHIRPLVSIYIPAYNCANYIQRQVDSALNQTVDLEVICNDGST 473
Db 421 VPIYRKLLPIEDSHIRPLVSIYIPAYNCANYIQRQVDSALNQTVDLEVICNDGST 480
Qy 474 DNTLEVINKLYGNPRVRIMSKPNGGIASASNAAYSFAGYYIGQLSDSDYLEPDAVELC 533
Db 481 DNTLEVINKLYGNPRVRIMSKPNGGIASASNAAYSFAGYYIGQLSDSDYLEPDAVELC 540
Qy 534 LKEFLKDKTLACVYTTNRVNPDSGLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 593
Db 541 LKEFLKDKTLACVYTTNRVNPDSGLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 600
Qy 594 TDGFENENIENAVDVMFLKSEVGKFKHLNKCYNRVLHGDNTSIKKLGIQKKNHFVVYN 653
Db 601 TDGFENENIENAVDVMFLKSEVGKFKHLNKCYNRVLHGDNTSIKKLGIQKKNHFVVYN 660
Qy 654 QSLNRQGINNYNDKFDLDSERKIFNKTAEYQEMDMKDLKLTQNKDKAIAVSIFYP 713
Db 661 QSLNRQGINNYNDKFDLDSERKIFNKTAEYQEMDMKDLKLTQNKDKAIAVSIFYP 720
Qy 714 NTLNGLVKKLNNIIEYNKNFVILHVDKNHLPDQIKKEILAFYHKHVNILLNNDISY 773
Db 721 NTLNGLVKKLNNIIEYNKNFVILHVDKNHLPDQIKKEILAFYHKHVNILLNNDISY 780
Qy 774 TSNRLIKTEAHLNLSNKLSQLNCEYIIFDNHDSLFVKNDISYAYKKYDVGWGNFSALTH 833
Db 781 TSNRLIKTEAHLNLSNKLSQLNCEYIIFDNHDSLFVKNDISYAYKKYDVGWGNFSALTH 840
```

```
Qy 834 DWIEKINAHPPFKKLIKTYFENDNDRSMNVKGSQGMFKMYALPHELLTIKKEIVITSQS 893
Db 841 DWIEKINAHPPFKKLIKTYFENDNDRSMNVKGSQGMFKMYALPHELLTIKKEIVITSQS 900
Qy 894 IDSVPEYNTEDIWFQPALILILEKKTGHVFNKTSITLTYPWERKLTQWNTQIOAKKGENI 953
Db 901 IDSVPEYNTEDIWFQPALILILEKKTGHVFNKTSITLTYPWERKLTQWNTQIOAKKGENI 960
Qy 954 PVNKKFIINSITL 965
Db 961 PVNKKFIINSITL 972

RESULT 6
US-09-469-200D-9
; Sequence 9, Application US/09469200D
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H.
; APPLICANT: DeAngelis, Paul
; APPLICANT: Kumari, Kshama
; TITLE OF INVENTION: Hyaluronan Synthase Gene and Uses Thereof
; FILE REFERENCE: 3554.011
; CURRENT APPLICATION NUMBER: US/09/469,200D
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: US 09/178,851
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: US 60/064,435
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 9
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Pastuerella Multocida
US-09-469-200D-9

Query Match      88.2%; Score 4490.5; DB 5; Length 972;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 843; Conservative 62; Mismatches 60; Indels 7; Gaps 2;

Qy 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGKIVFQIICKKEKLSN-----STNS 53
Db 1 MNTLSQAIKAYNSNDYQALKLFKSAETYGKIVFQIICKKEKLSAHPVNSAHLNVN 60
Qy 54 EDKNSVCDSSLDIATQALLSNVKKLTLSSEKNSLKNKWSITGKSENAEIRKVELVP 113
Db 61 EXEKNVCDSPDLIATQALLSNVKKLTLSXSEKNSLKNKWLITEKSENAEIRKVELVP 120
Qy 114 KDFPKDLVLAFLPDHVNDFTWYKNRKSLSGKPKVKNKIGLSIIPTFNRSLDITLACL 173
Db 121 KDFPKDLVLAFLPDHVNDFTWYKNRKSLSGKPKVKNKIGLSIIPTFNRSLDITLACL 180
Qy 174 VNQKTYPEVVDGSGKENLTTIVQYEQKLDITKYVRQKDYGYQLCAVRNLGLRTAKY 233
Db 181 VNQKTYPEVVDGSGQEXLLPLIXROYEXKLDITRYVRQKDYGYQACAAARXGLRLAKY 240
Qy 234 DFVSTLDCDMPAQOLWVHSYLTLELDDNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 293
Db 241 DFVSTLDCDMPAQOLWVHSYLTLELDDNDITIGPRKYVDTHNITAEQFLNDPYLIESLP 300
Qy 294 ETATNNPNSITSGNISLDRLEHFKTDNLRCLDPSERYFVAGNVAFSKWLNKVGWFD 353
Db 301 ETATNNPNSITSGNISLDRLEHFKTDNLRCLDPSERYFVAGNVAFSKWLNKVGWFD 360
Qy 354 EEFNHGGEDEVFGYRLFAKGCFFRVIDGMAIHQEPGKENETREAGKSTTLKIVREK 413
Db 361 EEFNHGGEDEVFGYRLFAKGCFFRVIDGMAIHQEPGKENETREAGKSTTLKIVREK 420
Qy 414 VPIYRKLLPIEDSHIRPLVSIYIPAYNCANYIQRQVDSALNQTVDLEVICNDGST 473
Db 421 VPIYRKLLPIEDSHIRPLVSIYIPAYNCANYIQRQVDSALNQTVDLEVICNDGST 480
Qy 474 DNTLEVINKLYGNPRVRIMSKPNGGIASASNAAYSFAGYYIGQLSDSDYLEPDAVELC 533
```

```
Db 481 DNTLEVNKLYGNNPRVIRMSKPGNGIASASNAASFAGYIIGQLSDDDYLEPDAVELC 540
QY 534 LKEFLDKDTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 593
Db 541 LKEFLDKDTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 600
QY 594 TDGFENENIENAVDYDMFLKLSVEGKFKHLNKCYNRVLHGDNTSIKKLGIOKKNHFFVVN 653
Db 601 TDGFENEKIENAVDYDMFLKLSVEGKFKHLNKCYNRVLHGDNTSIKKLGIOKKNHFFVVN 660
QY 654 QSLNRQGINYYNDKFDLDESRYIFNKTAEOEEMDLKDLKLTIONKDAKIAVSIYFP 713
Db 661 QSLNRQGIYYNDEFDLDESRYIFNKTAEOEEDILDKIKIIONKDAKIAVSIYFP 720
QY 714 NTLNGLVKLNNIIEYNKNFVILHVDKNHLPDIKKEILAFYHKHGVNILLNNDISYY 773
Db 721 NTLNGLVKLNNIIEYNKNFVILHVDKNHLPDIKKEILAFYHKHGVNILLNNDISYY 780
QY 774 TSRLIKTEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGWGFSAETH 833
Db 781 TSRLIKTEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGWGFSAETH 840
QY 834 DWIEKINAHPPFKLIKTYFNDNDRSMNVKGASQGMFKYALPHELLTIIKEVITSQOS 893
Db 841 DWIEKINAHPPFKLIKTYFNDNDRSMNVKGASQGMFKYALPHELLTIIKEVITSQOS 900
QY 894 IDSVPEYNTEDIWFOFALLILEKKTGHVFNKTSLTTPWPERKLOWTNEQIOAKKGENI 953
Db 901 IDSVPEYNTEDIWFOFALLILEKKTGHVFNKTSLTTPWPERKLOWTNEQIOAKKGENI 960
QY 954 PVNKFINSITL 965
Db 961 PVNKFINSITL 972

RESULT 7
US-10-309-560-8
; Sequence 8, Application US/10309560
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H
; APPLICANT: KUMAR, KSHAMA
; TITLE OF INVENTION: HYALURONAN SYNTHASES AND METHODS OF MAKING AND USING SAME
; FILE REFERENCE: 35541.082
; CURRENT APPLICATION NUMBER: US/10/309,560
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: 60/336,105
; PRIOR FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 972
; TYPE: PR1
; ORGANISM: Pasteurella multocida
US-10-309-560-8

Query Match 88.2%; Score 4486.5; DB 6; Length 972;
Best Local Similarity 86.6%; Pred. No. 0;
Matches 842; Conservative 63; Mismatches 60; Indels 7; Gaps 2;

QY 1 MNTLSQAIKATNSNDYELALKLFPEKSAETYGKRIKVEFOIRCKEKL----STNS---YVS 53
Db 1 MNTLSQAIKATNSNDYELALKLFPEKSAETYGKRIKVEFOIRCKEKL---SHPVNSAHLV 60
QY 54 EDKKNVCSDDLDTATQLLSNVKLTLSSEKNSLKNKKSITGKKSSENAEIRKVELVP 113
Db 61 KEKVNVCSDDLDTATQLLSNVKLVLSDSSEKNTLKNKWLTLTKKSENAEVRVALVP 120
QY 114 KDFPKDLVLPDHPNDFTWYKKNKSLGKPKVKNKIGLSIIITFNRSRILDTLACL 173
Db 121 KDFPKDLVLPDHPNDFTWYKKNKSLGKPKVKNKIGLSIIITFNRSRILDTLACL 180
QY 174 VNQKTNYPFEVVVADGSGKENLTTIVQKYEOKLDIKYVRQKDYQOLCAVRNIGLRTAKY 233
```

```
Db 181 VNQKTHYPFEVITVDGSGQEDLSPIIRQYENKLDIRVROKDNQFQSAARNMGLRLAKY 240
QY 234 DFFVSLDCDMAPQOLWVHSLYTELLEDNDIVLIGPRXYVDTHNITAEQFLNDPVLIESLP 293
Db 241 DFTGLDCDMAPNPLWVHSLYAELEDDDLTIIGPRXYIDTQHIDPKDFLNNASLESPL 300
QY 294 ETATNNPSTTSKGNISLDWRLEHFKKTDNLRLCDSPRFVAGNVAFSKELWLNKVGWFD 353
Db 301 EVKTNNVAAKGEQTVSLDWRLEQFEXTENLRLSDSFPRFAAGNVAFKAKWLNKSGFFD 360
QY 354 EEFNHWGGEDEVEFYRLFAKGCFFRVIDGGMAIHQEPGPKENETEREAGKSIITLKIYVEK 413
Db 361 EEFNHWGGEDEVEFYRLFRYGSFFKTDIGIMAYHQEPGPKENETEREAGKNTITDIMREK 420
QY 414 VPIYRKLLPIEDSHIRIPLVSIYIPAYNCANYIQRCVDSALNQTVVVDLVCVNCNDGST 473
Db 421 VPIYRKLLPIEDSHIRNPLVSIYIPAYNCANYIQRCVDSALNQTVVVDLVCVNCNDGST 480
QY 474 DNTLEVINKLYGNNPRVIRMSKPGNGIASASNAASFAGYIIGQLSDDDYLEPDAVELC 533
Db 481 DNTLEVINKLYGNNPRVIRMSKPGNGIASASNAASFAGYIIGQLSDDDYLEPDAVELC 540
QY 534 LKEFLDKDTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 593
Db 541 LKEFLDKDTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 600
QY 594 TDGFENENIENAVDYDMFLKLSVEGKFKHLNKCYNRVLHGDNTSIKKLGIOKKNHFFVVN 653
Db 601 TDGFENEKIENAVDYDMFLKLSVEGKFKHLNKCYNRVLHGDNTSIKKLGIOKKNHFFVVN 660
QY 654 QSLNRQGINYYNDKFDLDESRYIFNKTAEOEEMDLKDLKLTIONKDAKIAVSIYFP 713
Db 661 QSLNRQGIYYNDEFDLDESRYIFNKTAEOEEDILDKIKIIONKDAKIAVSIYFP 720
QY 714 NTLNGLVKLNNIIEYNKNFVILHVDKNHLPDIKKEILAFYHKHGVNILLNNDISYY 773
Db 721 NTLNGLVKLNNIIEYNKNFVILHVDKNHLPDIKKEILAFYHKHGVNILLNNDISYY 780
QY 774 TSRLIKTEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGWGFSAETH 833
Db 781 TSRLIKTEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGWGFSAETH 840
QY 834 DWIEKINAHPPFKLIKTYFNDNDRSMNVKGASQGMFKYALPHELLTIIKEVITSQOS 893
Db 841 DWIEKINAHPPFKLIKTYFNDNDRSMNVKGASQGMFKYALPHELLTIIKEVITSQOS 900
QY 894 IDSVPEYNTEDIWFOFALLILEKKTGHVFNKTSLTTPWPERKLOWTNEQIOAKKGENI 953
Db 901 IDSVPEYNTEDIWFOFALLILEKKTGHVFNKTSLTTPWPERKLOWTNEQIOAKKGENI 960
QY 954 PVNKFINSITL 965
Db 961 PVNKFINSITL 972

RESULT 8
US-10-217-613-1
; Sequence 1, Application US/10217613
; GENERAL INFORMATION:
; APPLICANT: DEANGELIS, PAUL
; TITLE OF INVENTION: DNA ENCODING HYALURONAN SYNTHASE FROM PASTURELLA MULTOCIDA AND
; FILE REFERENCE: 35541.081
; CURRENT APPLICATION NUMBER: US/10/217,613
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/283,402
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 9
```


; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 703

; TYPE: PRT

; ORGANISM: Pasteurella multocida

US-10-217-613-1

Query Match 61.0%; Score 3102.5; DB 6; Length 703;

Best Local Similarity 82.6%; Pred. No. 1.5e-287;

Matches 581; Conservative 57; Mismatches 58; Indels 7; Gaps 2;

Qy 1 MNTLSQAIKAYNSNDYELALKLFPEKSAETYGKRIKVEFOIICKCKEL-----STNS---YVS 53

Db 1 MNTLSQAIKAYNSNDYQALALKLFPEKSAEYGRKIVFEQITCKCKELSAHPSVNSAHLNVN 60

Qy 54 EDKNSVCSSLDIATQALLSNVKKLTLSSEKNSLKNKWSITGKKSENAEIRKVELVP 113

Db 61 KEKYNVCDSPDLIATQALLSNVKKLVLSDEKNTLNKWKLLTEKXSENAEYRAVALVP 120

Qy 114 KDFPKDLVLAHPDVNDFTWYKNRKSIGIKPVNKNIGLSIIIPFNRSRILDTITLACL 173

Db 121 KDFPKDLVLAHPDVNDFTWYKKRKRGLGKPEHQHVGLSIIVTFNRPAILLSITLACL 180

Qy 174 VNQKNTPEFVVDGSKENLLTVQKYEQKLDIKYVRQKDYGYQLCAVRNGLRTAKY 233

Db 181 VNQKTHYPEVIVTDGSOEDLSPIIROYENKLDIRYVRQKDNQGFQASAARNMGLRLAKY 240

Qy 234 DFVSTLDCDMPQOVLVHSHYLTLELDDNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 293

Db 241 DFIGLLDCDMPNPLVHSHYVAELLEDLTIIGPRKYIDTQHIDPKDFLNNASLESLEP 300

Qy 294 ETATNNPNSITSKGNISLDWRLEHFKKTDNRLCLDPSPFYFVAGNVAFSKWLNKVGWFD 353

Db 301 EVKTNNSVAAKGEVTVSLDWRLEQFEKTENLRSLSDSPFRFFAAGNVAFKAKWLNKSGFFD 360

Qy 354 EEFNHMGDEVEFGYRLFAKCGFFRVIDGMAHQEPGKNETEREAGKSTTLKIVKEK 413

Db 361 EEFNHMGDEVEFGYRLFRYGSFEKTIDGIMAYHQEPGKNETEREAGKNTLIDIMREK 420

Qy 414 VPYIRKLLPTEDSHRIPLVSIYPAYNCANYTORCVDSALNQTVDLEVCICNDGST 473

Db 421 VPYIRKLLPTEDSHINRVPLVSIYPAYNCANYIQRVDSALNQTVDLEVCICNDGST 480

Qy 474 DNTLEVINKLYGNPRVIMSKPNGGSIASNAAVSFAGYYIGQLSDDDYLEPDAVELC 533

Db 481 DNTLEVINKLYGNPRVIMSKPNGGSIASNAAVSFAGYYIGQLSDDDYLEPDAVELC 540

Qy 534 LKEFLKDKTLACVYTTNRVNPDSGLIANGYNWPEFSREKLTMTAMIAHFRMFTTIRAWHL 593

Db 541 LKEFLKDKTLACVYTTNRVNPDSGLIANGYNWPEFSREKLTMTAMIAHFRMFTTIRAWHL 600

Qy 594 TDGFNEKNIENAVDYMFLKLVSEVGKFKHLNKCYNRVLRHGDNSTSKLGIQKKNHFVVVN 653

Db 601 TDGFNEKNIENAVDYMFLKLVSEVGKFKHLNKCYNRVLRHGDNSTSKLGIQKKNHFVVVN 660

Qy 654 QSLNRQGINNYNDKFDLDESRYIFNKTAEQEEMDLKDL 696

Db 661 QSLNRQGIYYNYDEDFDLDESRYIFNKTAEQEEDILDKI 703

RESULT 9

US-10-217-613-7

; Sequence 7, Application US/10217613

; GENERAL INFORMATION:

; APPLICANT: DEANGELIS, PAUL

; TITLE OF INVENTION: DNA ENCODING HYALURONAN SYNTHASE FROM PASTURELLA MULTOCIDA AND US

; FILE OF INVENTION: THEREOF

; FILE REFERENCE: 35541.081

; CURRENT APPLICATION NUMBER: US/10/217,613

; CURRENT FILING DATE: 2002-08-12

; PRIOR APPLICATION NUMBER: 09/283,402

; PRIOR FILING DATE: 1999-04-01

; PRIOR APPLICATION NUMBER: 60/080,414

; PRIOR FILING DATE: 1998-04-02

; PRIOR APPLICATION NUMBER: 09/178,851

; PRIOR FILING DATE: 1998-10-26

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 7

; LENGTH: 703

; TYPE: PRT

; ORGANISM: Pasteurella multocida

US-10-217-613-7

Query Match 61.0%; Score 3102.5; DB 6; Length 703;

Best Local Similarity 82.6%; Pred. No. 1.5e-287;

Matches 581; Conservative 57; Mismatches 58; Indels 7; Gaps 2;

Qy 1 MNTLSQAIKAYNSNDYELALKLFPEKSAETYGKRIKVEFOIICKCKEL-----STNS---YVS 53

Db 1 MNTLSQAIKAYNSNDYQALALKLFPEKSAEYGRKIVFEQITCKCKELSAHPSVNSAHLNVN 60

Qy 54 EDKNSVCSSLDIATQALLSNVKKLTLSSEKNSLKNKWSITGKKSENAEIRKVELVP 113

Db 61 KEKYNVCDSPDLIATQALLSNVKKLVLSDEKNTLNKWKLLTEKXSENAEYRAVALVP 120

Qy 114 KDFPKDLVLAHPDVNDFTWYKNRKSIGIKPVNKNIGLSIIIPFNRSRILDTITLACL 173

Db 121 KDFPKDLVLAHPDVNDFTWYKKRKRGLGKPEHQHVGLSIIVTFNRPAILLSITLACL 180

Qy 174 VNQKNTPEFVVDGSKENLLTVQKYEQKLDIKYVRQKDYGYQLCAVRNGLRTAKY 233

Db 181 VNQKTHYPEVIVTDGSOEDLSPIIROYENKLDIRYVRQKDNQGFQASAARNMGLRLAKY 240

Qy 234 DFVSTLDCDMPQOVLVHSHYLTLELDDNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 293

Db 241 DFIGLLDCDMPNPLVHSHYVAELLEDLTIIGPRKYIDTQHIDPKDFLNNASLESLEP 300

Qy 294 ETATNNPNSITSKGNISLDWRLEHFKKTDNRLCLDPSPFYFVAGNVAFSKWLNKVGWFD 353

Db 301 EVKTNNSVAAKGEVTVSLDWRLEQFEKTENLRSLSDSPFRFFAAGNVAFKAKWLNKSGFFD 360

Qy 354 EEFNHMGDEVEFGYRLFAKCGFFRVIDGMAHQEPGKNETEREAGKSTTLKIVKEK 413

Db 361 EEFNHMGDEVEFGYRLFRYGSFEKTIDGIMAYHQEPGKNETEREAGKNTLIDIMREK 420

Qy 414 VPYIRKLLPTEDSHRIPLVSIYPAYNCANYIQRVDSALNQTVDLEVCICNDGST 473

Db 421 VPYIRKLLPTEDSHINRVPLVSIYPAYNCANYIQRVDSALNQTVDLEVCICNDGST 480

Qy 474 DNTLEVINKLYGNPRVIMSKPNGGSIASNAAVSFAGYYIGQLSDDDYLEPDAVELC 533

Db 481 DNTLEVINKLYGNPRVIMSKPNGGSIASNAAVSFAGYYIGQLSDDDYLEPDAVELC 540

Qy 534 LKEFLKDKTLACVYTTNRVNPDSGLIANGYNWPEFSREKLTMTAMIAHFRMFTTIRAWHL 593

Db 541 LKEFLKDKTLACVYTTNRVNPDSGLIANGYNWPEFSREKLTMTAMIAHFRMFTTIRAWHL 600

Qy 594 TDGFNEKNIENAVDYMFLKLVSEVGKFKHLNKCYNRVLRHGDNSTSKLGIQKKNHFVVVN 653

Db 601 TDGFNEKNIENAVDYMFLKLVSEVGKFKHLNKCYNRVLRHGDNSTSKLGIQKKNHFVVVN 660

Qy 654 QSLNRQGINNYNDKFDLDESRYIFNKTAEQEEMDLKDL 696

Db 661 QSLNRQGIYYNYDEDFDLDESRYIFNKTAEQEEDILDKI 703

RESULT 10

US-10-216-289-2

; Sequence 2, Application US/10216289

; GENERAL INFORMATION:

; APPLICANT: NINOMIYA, TOSHIO

; APPLICANT: SUGIURA, NOBUO

; APPLICANT: KIMATA, KOJI

; TITLE OF INVENTION: CHONDROITIN POLYMERASE AND DNA ENCODING THE SAME

; FILE REFERENCE: 226882USO

```
; CURRENT APPLICATION NUMBER: US/10/216,289
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: JP 2001-244685
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: JP 2001-324127
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: JP 2002-103136
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-216-289-2
```

```
Query Match 44.3%; Score 2256.5; DB 6; Length 686;
Best Local Similarity 61.3%; Pred. No. 9.8e-207;
Matches 419; Conservative 112; Mismatches 148; Indels 5; Gaps 3;

QY 1 MNTLSQAIKAYNSNDYELALKLFKSAETVGRKIVEFQIIKCKEKLSTNSYVSEDKKNV 60
Db 1 MSTLNQAINLYKNKNYQALSLFEKVAEYDVSWEANIKLCQTALNLSVEVDKLNKAV 60

QY 61 CDSLSDIATQLLSNVKKLTLSSEKSNLKNWKSITGKKSSENAEIRKVELVPKDFPKDL 120
Db 61 ID--IDAATKIMCSNAKAISLNEVEKNEIISKYREITAKKSERAEKVEVEPIPLDWPSDL 118

QY 121 VLAPLPHVNDFTWYKNRKKSLGKIPVKNKI--GLSIITPTFRNSRLDITLACLVNQKT 178
Db 119 TLPPLPESTNDYVW--AGKRELDDYPRKQLIIDGLSIVIPYTNRAKILAITLACLQNKT 177

QY 179 NYPEFVVVADGSGKENLLTIVQYEQKLDIKYVRQKDYGYQLCAVRNLGRTAKYDFVSI 238
Db 178 IYDEYIVVADGSGKENIEEIVREFESLLNKKYVRQKDYGYQLCAVRNLGLRAAKYNYVAI 237

QY 239 LDCDMPAQLWVHSYLTLELLEDNDIVLIGRPKYVDTHNITAEQFLNDPYLIESLPETATN 298
Db 238 LDCDMPANPLWQSYMELLAVDNDVALIGRPKYIDTSKHTYLDLFSQKSLINEIPEITN 297

QY 299 NNPSITSGKNISLDWRLHFKKTDNLRLCDSPPRYFVAGNVAFSKWLNKVGWDFDEEFNH 358
Db 298 NQVAGVQKNSVDWRLHFKKTDNLRLCNTPPRFSSGNVAFKAKWLFRAFWDFDEETH 357

QY 359 WGEDVEFGYRLFAKGFRRVIDGGMALHOEPPGKENETEREAGKSTTLKIVKEKVPYIY 418
Db 358 WGEDNEFGYRLYREGCYFRSVEGAMAYHOEPPGKENETDRAAGKNITVQLLOQKVPYFY 417

QY 419 RKLPIEDSHIHRPLYSIYPAYNCANYIQRCDVSALNOTVVDLEVICINDGSTDNTLE 478
Db 478 ILQEHYANHPVRPFIQKNGIGASANTAVRLCRGFFYIGQLSDDDLEPDAVELCLDEFR 537

QY 539 KDKTLACVYTTNRNVPDGLIANGYNWPEFSREKLTMTAMIAHFRFTTIRAWHLTDGFN 598
Db 538 KDLSLACVYTTNRNIDREGNLISNGYNWPIYSREKLTSAMICHHFRTTARANNLTGEGN 597

QY 599 ENTENAVDYDMFLKLSVGVKFKHLNKICYNRVLHGDNWTSIKKLGIOKKHFFVYVYNQSLNR 658
Db 598 ESISNADYDYMFLKLSVGVGPFKINKICYNRVLHGENTSIIKLDIQENHFYVYNESLSR 657

QY 659 QGINNYNDKFDLDESRYKIFNK 682
Db 658 LGIKKYKYSPLTNLNECRKYTWEK 681
```

```
RESULT 11
US-10-216-289-4
; Sequence 4, Application US/10216289
; GENERAL INFORMATION:
```

```
; APPLICANT: NINOMIYA, TOSHIO
; APPLICANT: SUGIURA, NOBUO
; APPLICANT: KIMATA, KOJI
; TITLE OF INVENTION: CHONDROITIN POLYMERASE AND DNA ENCODING THE SAME
; FILE REFERENCE: 226882USO
; CURRENT APPLICATION NUMBER: US/10/216,289
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: JP 2001-244685
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: JP 2001-324127
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: JP 2002-103136
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-216-289-4

Query Match 44.3%; Score 2256.5; DB 6; Length 686;
Best Local Similarity 61.3%; Pred. No. 9.8e-207;
Matches 419; Conservative 112; Mismatches 148; Indels 5; Gaps 3;

QY 1 MNTLSQAIKAYNSNDYELALKLFKSAETVGRKIVEFQIIKCKEKLSTNSYVSEDKKNV 60
Db 1 MSTLNQAINLYKNKNYQALSLFEKVAEYDVSWEANIKLCQTALNLSVEVDKLNKAV 60

QY 61 CDSLSDIATQLLSNVKKLTLSSEKSNLKNWKSITGKKSSENAEIRKVELVPKDFPKDL 120
Db 61 ID--IDAATKIMCSNAKAISLNEVEKNEIISKYREITAKKSERAEKVEVEPIPLDWPSDL 118

QY 121 VLAPLPHVNDFTWYKNRKKSLGKIPVKNKI--GLSIITPTFRNSRLDITLACLVNQKT 178
Db 119 TLPPLPESTNDYVW--AGKRELDDYPRKQLIIDGLSIVIPYTNRAKILAITLACLQNKT 177

QY 179 NYPEFVVVADGSGKENLLTIVQYEQKLDIKYVRQKDYGYQLCAVRNLGRTAKYDFVSI 238
Db 178 IYDEYIVVADGSGKENIEEIVREFESLLNKKYVRQKDYGYQLCAVRNLGLRAAKYNYVAI 237

QY 239 LDCDMPAQLWVHSYLTLELLEDNDIVLIGRPKYVDTHNITAEQFLNDPYLIESLPETATN 298
Db 238 LDCDMPANPLWQSYMELLAVDNDVALIGRPKYIDTSKHTYLDLFSQKSLINEIPEITN 297

QY 299 NNPSITSGKNISLDWRLHFKKTDNLRLCDSPPRYFVAGNVAFSKWLNKVGWDFDEEFNH 358
Db 298 NQVAGVQKNSVDWRLHFKKTDNLRLCNTPPRFSSGNVAFKAKWLFRAFWDFDEETH 357

QY 359 WGEDVEFGYRLFAKGFRRVIDGGMALHOEPPGKENETEREAGKSTTLKIVKEKVPYIY 418
Db 358 WGEDNEFGYRLYREGCYFRSVEGAMAYHOEPPGKENETDRAAGKNITVQLLOQKVPYFY 417

QY 419 RKLPIEDSHIHRPLYSIYPAYNCANYIQRCDVSALNOTVVDLEVICINDGSTDNTLE 478
Db 478 ILQEHYANHPVRPFIQKNGIGASANTAVRLCRGFFYIGQLSDDDLEPDAVELCLDEFR 537

QY 539 KDKTLACVYTTNRNVPDGLIANGYNWPEFSREKLTMTAMIAHFRFTTIRAWHLTDGFN 598
Db 538 KDLSLACVYTTNRNIDREGNLISNGYNWPIYSREKLTSAMICHHFRTTARANNLTGEGN 597

QY 599 ENTENAVDYDMFLKLSVGVKFKHLNKICYNRVLHGDNWTSIKKLGIOKKHFFVYVYNQSLNR 658
Db 598 ESISNADYDYMFLKLSVGVGPFKINKICYNRVLHGENTSIIKLDIQENHFYVYNESLSR 657

QY 659 QGINNYNDKFDLDESRYKIFNK 682
Db 658 LGIKKYKYSPLTNLNECRKYTWEK 681
```


; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5087
; LENGTH: 534
; TYPE: PRF
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-5087

Query Match 5.2%; Score 262.5; DB 5; Length 534;
Best Local Similarity 23.0%; Pred. No. 1.9e-16;
Matches 133; Conservative 100; Mismatches 234; Indels 111; Gaps 22;

Qy 260 DNDIVLAGPRKYVD--THNITAEQFLNDPVLIESLPETATNNPSITSKGNISLDWR--- 314
Db 15 NEDIKIFDSIYRDKTNNLTITGALDTIKES--PTFTINNENQYSAYNIQVRLREDV 72
Qy 315 -----LEHFKTDNLRCDSPERYFVAGNVAFSEKWLNK-----VQWF 352
Db 73 NQIYQTEPAIEAGFVWTLLEGIKQKKVL-----PFHQSSAHVITVDFPLNKKYPVPGTE 127
Qy 353 DEEFNHGGEDVEGYRLFAGKCFRVIDGGMATHQEPGKNETEREPAGKSIILKIVKE 412
Db 128 DKVTRLW-----IKARKGPKYMAKNGIS--HTQRAKIEKLRLQA--SYPNWLARN 174
Qy 413 KVPYIRKLLPIEDSHIRIPLVSIYIPAYNC-ANYIQRVCDSALNOTVVDLEVCICNDG 471
Db 175 EVLDI--EAMTQETATPHYQPKISAMPVYNVEEKWLRLCIDSILNQVYTNWELCMADDA 232
Qy 472 STD-NLDEVINKLNNPRVRIM-SKPNGGIASNAVAFAKGYIGQLSDDDYLEPDA 529
Db 233 STDNPKKILTEYQQLDERIRVFRREQNGHISEATNSALATATGEFVALLDNDDELAINA 292
Qy 530 VELCLKEFLDKTACVYTTNRNVNPDGSLIANGYNWPEFSREKLTMTAMIAHFFRMFTIR 589
Db 293 FYEVKVLNENPELDLIYSDDEKIDMDGNRSDPAFK-FWSPDLLLTGTYISHLGVYRS 351
Qy 590 AWHLTDFGNENIENAVDYMFLKSE---VGKFKHLNLCYNRVLHGDNTSIKK--LGIQ 644
Db 352 ILEETGFRKGYEGSQDYDLVLRFTKTKERIKHPIKVLXYWRMLPTSTAVDQGSKGYA 411
Qy 645 KKHVFWVNVQSLNRQGINYN-----YDKFDLDESRYIFNKTAEYQEMDMKLDL 696
Db 412 FEAGLRAYODALVRGNGHATHGAANGLYDYIDI-ESEK----- 451
Qy 697 KLIQNKDAKIAVSIFYPNTLNG---LVKLNNIIEYNKNIFVIIHLVDKNHLP----- 747
Db 452 -----LVSIILP-TKNYKDVQRCVSSIIETTYQNYQVYIIMADNGSTDPRKHELY 500
Qy 748 -DIKKEILAFYHKHQNILLNNDISYYTSNRLIKTEAH 784
Db 501 AEFEOQLPGRFVESIDIPFN---FSTINNRAAKKAH 534

RESULT 15
US-09-134-000C-6392
; Sequence 6392, Application US/09134000C
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6392
; LENGTH: 330
; TYPE: PRF
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-6392

Query Match 4.9%; Score 248.5; DB 5; Length 330;
Best Local Similarity 24.9%; Pred. No. 1.8e-15;
Matches 85; Conservative 63; Mismatches 125; Indels 69; Gaps 14;

Qy 432 IPLVSIYIPAYNCANYIQRVCDSALNOTVVDLEVCICNDGSTDNTLEVINLKGNNPRVR 491
Db 7 MPKISIIIVPVYNVEKYLEKCVRSIIAQFTDFELIYDDGSPDSGAMCDQFAEQDQVRK 66
Qy 492 IMSPNNGGIASASNAVAFAKGYIGQLSDDDYLEPDAVELCLKEFLK---DKTLACYVT 548
Db 67 VIHRENGGLSDARNAGIELATGEYLGFDVSDDIYADDMVELLYTNIVKEDADLSICGIYD 126
Qy 549 TNRNVNPDGSLIANGYNWPEFSREKLTMTAMIAHFFRMFTIRAWHLTDGFENENIENAVDYD 608
Db 127 VYEGKEPIVKSIIQ-----TFSREALLILQGN--IISVHA----- 162
Qy 609 MFLKLSYGVKFKHLNLCYNRVLHGDNTSIKKIGIQKKNHFWVNVQSLNRQGIN----- 662
Db 163 -----VNKL-YKRKLFADLRYPK--GYHEDSFIIVDLLSECQKVSIDSTOK 206
Qy 663 YNYNDKFDLDESRYIFENKTAEYQEMDMKLDKLIQNKDAKIAVS---IFYPNTL-- 716
Db 207 YYYHHRMGSII--NTEFSDKQFETAEWEK-NELKL-KGKGAVIEAAAHQRCVFANFLVL 262
Qy 717 -----NGLVKK-LNNIIEYNKNIFVIIHLVDKNHLPDIIK 751
Db 263 DKILISNAPKKKTKQIVRYLRFENFIFIM---KNKVFTKSRK 301

Search completed: January 4, 2003, 02:45:48
Job time : 565 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2003, 02:18:59 ; Search time 69 seconds
(without alignments)
1344.489 Million cell updates/sec

Title: US-09-842-484A-2
Perfect score: 5089
Sequence: 1 MNTLSQAIKAYNSNDYELAL.....SAKKGNIPIVKNFIINSITL 965

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4486.5	88.2	972	2 T09595	glucuronosyltransf
2	294	5.8	604	2 E97757	hypothetical prote
3	292.5	5.7	327	2 AB1211	glycosyltransferas
4	268.5	5.3	696	2 D95206	hypothetical prote
5	264.5	5.2	323	2 AD2189	hypothetical prote
6	261.5	5.1	303	2 A84114	glycosyltransferas
7	258.5	5.1	333	2 B97168	glycosyltransferas
8	256	5.0	321	2 AG2188	hypothetical prote
9	254.5	5.0	324	2 AB2190	hypothetical prote
10	251	4.9	298	2 B75096	glycosyl transfera
11	241	4.7	333	2 AH2026	hypothetical prote
12	239	4.7	340	2 T44330	glycosyl transfera
13	238	4.7	306	2 T50038	beta-1,3-N-acetylgl
14	238	4.7	316	2 AE2189	hypothetical prote
15	237.5	4.7	257	2 E84107	telchuronic acid b
16	237	4.7	344	2 A70037	capsular polysacch
17	235.5	4.6	318	1 E71690	minor teichoic aci
18	235	4.6	318	2 AH2189	hypothetical prote
19	234.5	4.6	336	2 A97168	glycosyltransferas
20	234	4.6	322	2 T44647	glycosyl transfera
21	233.5	4.6	1013	2 AE1876	hypothetical prote
22	232	4.6	333	2 H97167	glycosyltransferas
23	229.5	4.5	349	2 D81027	lacto-N-neotetraos
24	229.5	4.5	776	2 F81289	probable sugar tra
25	229	4.5	389	2 E81318	probable galactosy
26	228	4.5	392	2 H69814	hypothetical prote
27	227.5	4.5	333	1 S70813	glycosyl transfera
28	224.5	4.4	323	1 H64130	glycosyl transfera
29	223.5	4.4	338	2 E91190	probable regulator

30	223.5	4.4	338	2 F86037	probable regulator
31	223	4.4	318	2 T50039	beta-1,4-galactosy
32	221	4.3	330	2 AH2188	hypothetical prote
33	221	4.3	334	1 G71153	hypothetical prote
34	221	4.3	338	2 E97083	glycosyltransferas
35	220.5	4.3	445	2 B81289	probable sugar tra
36	220	4.3	329	2 AG0023	probable glycosyl
37	218.5	4.3	318	2 AG2189	hypothetical prote
38	217.5	4.3	344	1 Q3ECTH	hypothetical 40.5K
39	217	4.3	367	2 G95948	probable glycosylt
40	216	4.2	390	2 C81318	probable galactosy
41	215.5	4.2	346	2 H81970	lacto-N-neotetraos
42	214.5	4.2	301	2 F95205	glycosyl transfera
43	214.5	4.2	515	2 B81318	probable two-domai
44	214	4.2	250	2 A64099	glycosyl transfera
45	214	4.2	313	2 AI2404	hypothetical prote

ALIGNMENTS

RESULT 1
T09595

glucuronosyltransferase (EC 2.4.1.17) - Pasteurella multocida

N:Alternate names: hyaluronan synthase

C:Species: Pasteurella multocida

C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000

C:Accession: T09595

R:DeAngelis, P.L.; Jing, W.; Drake, R.R.; Achyuthan, A.M.

J. Biol. Chem. 273, 8454-8458, 1998

A:Title: Identification and molecular cloning of a unique hyaluronan synthase from

A:Reference number: Z16757; MUID:98192645; PMID:9525958

A:Accession: T09595

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-972 <DEA>

A:Cross-references: EMBL:AF036004; NID:G3043922; PIDN:AAC38318.1; PID:G3043923

A:Experimental source: strain P-1059; ATCC 15742

C:Genetics:

A:Gene: HAS

C:Function:

A:Description: polymerizes hyaluronan (HA, hyaluronate, hyaluronic acid) polysacchar
C:Keywords: capsule synthesis; glycosyltransferase; hexosyltransferase; hyaluronic a

Query Match 88.2%; Score 4486.5; DB 2; Length 972;
Best Local Similarity 86.6%; Pred. No. 1.2e-238;
Matches 842; Conservative 63; Mismatches 60; Indels 7; Gaps 2;

QY 1 MNTLSQAIKAYNSNDYELALFKFEKSAETGYGRKIVEFQIIKCKEKL----STNS---YVS 53

Db 1 MNTLSQAIKAYNSNDYQLALKFKFEKSAEYGRKIVEFQITKQEKLSAHPVNSAHLNVN 60

QY 54 EDKNSVCDSSLDIATOLLNSNVKLLTSESEKNSLNKWKSTGKKSNAEIRKVELVP 113

Db 61 KEKVNVVCDSPDLDTATOLLNSNVKLLVSDSEKNTLNKWLKTEKKSNAEYRAVALVP 120

QY 114 KDPKDLVLAPLDPHVNDFTWYKNRKSGLGTPKNKNGLSIIITPTNRSRLDITLACL 173

Db 121 KDPKDLVLAPLDPHVNDFTWYKNRKRRLGKLPKHQHVGLSIIIVTTNRPAILSIITLACL 180

QY 174 VNOKTNPFFVWVADGSKENLLITVQYEQKLDIKVYRQKDYGYQLCAVRLGLRTAKY 233

Db 181 VNOKTHYPFVIVTDDGSEDLSPIIQYENKLDIRVYRQKNGFQSAARNMGLRLAKY 240

QY 234 DFVSLDCDMAPQOLWVHSYLTLELNDIVLIGPRKYVDTHNTATQOFLNDPVLIESLP 293

Db 241 DFGLLLDCDMAPNPLWVHSYVAELLEDLTIIGPRKYIDTQHDPKDFLNNSLLESPL 300

QY 294 ETATNNPSTSGNISLDRLEHFKKTDNLRLCDSPFRFVAGNVAFSKENLKNVGF 353

Db 301 EVTNNVAAGEGTGVSLDRLEHFKKTDNLRLCDSPFRFVAGNVAFSKENLKNVGF 360

QY 354 EEFNHWGGEVDFGYRLFAKGCFFRVIDGGMATHQEPGKNETEREAGKSITLKVKEK 413

||||| 361 EEFHWGGEVFGRLPRYGSFFKIDGIMAYHQEPCKENETDREAGKNTLIDIMEK 420
QY 414 VPIYRKLLPIEDSHIRPLVSIYIPAYNCANYIQRCDVSALNQTVVDLEVCINDGST 473
Db 421 VPIYRKLLPIEDSHIRPLVSIYIPAYNCANYIQRCDVSALNQTVVDLEVCINDGST 480
QY 474 DNTLEVINKLGNNPRVIMSKPNGGSIASASNAVSPAKYIYGQDSDDDYLEPAVELC 533
Db 481 DNTLEVINKLGNNPRVIMSKPNGGSIASASNAVSPAKYIYGQDSDDDYLEPAVELC 540
QY 534 LKEFLKDKTCLACVYTTNNRNPVDPGSLIANGYNWPEFSREKLTAMIAHFRFTIRAWHL 593
Db 541 LKEFLKDKTCLACVYTTNNRNPVDPGSLIANGYNWPEFSREKLTAMIAHFRFTIRAWHL 600
QY 594 TDGFNENIENAVDYMFLKSEVGFKHLNKICYNRVLHGDNTSKLGIQKKNHFVVYN 653
Db 601 TDGFNEKIENAVDYMFLKSEVGFKHLNKICYNRVLHGDNTSKLGIQKKNHFVVYN 660
QY 654 QSLNRQGINNYNDFDLDDESKYIFNKTAQYQEMDKDLKLIQNKDAKIAVSIFYP 713
Db 661 QSLNRQGINNYNDFDLDDESKYIFNKTAQYQEIIDLKIKIIONKDAKIAVSIFYP 720
QY 714 NTLNGLVKKLNNIEYKNKFIIVLHVDKNHLLTPDKKEILAFYKHQVNNILLNNDISY 773
Db 721 NTLNGLVKKLNNIEYKNKFIIVLHVDKNHLLTPDKKEILAFYKHQVNNILLNNDISY 780
QY 774 TSNRLIKTEAHSNKLNSQLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGWGNFSALTH 833
Db 781 TSNRLIKTEAHSNKLNSQLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGWGNFSALTH 840
QY 834 DWIEKINAPPEKLIKTYTFNDNLRSMNVKASQGMFKYALPHELLTIKEVITSCQS 893
Db 841 DWIEKINAPPEKLIKTYTFNDNLRSMNVKASQGMFTYALAHELLTIKEVITSCQS 900
QY 894 IDSVPYNTEDIWFOFALLILEKKGHVFNKYSTLTYPWERKLTQNTNIOQSAKGENI 953
Db 901 IDSVPYNTEDIWFOFALLILEKKGHVFNKYSTLTYPWERKLTQNTNIOQSAKGENI 960
QY 954 PVNKFIIINSITL 965
Db 961 PVNKFIIINSITL 972
RESULT 2
E97757
hypothetical protein RC0461 [imported] - Rickettsia conorii (strain Malish 7)
C:Species: Rickettsia conorii
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
C:Accession: E97757
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R
Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: E97757
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-604 <KUR>
A:Cross-references: GB:AE006914; PIDN:AAL02999.1; PID:g15619533; GSPDB:GN00173
C:Genetics:
A:Gene: RC0461
Query Match 5.8%; Score 294; DB 2; Length 604;
Best Local Similarity 24.8%; Pred. No. 1.le-08;
Matches 124; Conservative 71; Mismatches 213; Indels 92; Gaps 19;
QY 419 RKLPIEDSHIRPLVSIYIPAYNCANYIQRCDVSALNQTVVDLEVCINDGSTDN-TL 477
Db 4 RKDKPLNIYH---TLVSIIPYINGANYMKEAINSALQTYKNIEIIVVNDGSKDNGET 59
QY 478 EVTNKLGNNPRVIMSKPNGGSIASASNAVSPAKYIYGQDSDDDYLEPAVE----LC 533
Db 60 ERVALSYGD--KIRYFYKENGCGGCSALNYGIKNQGEYFWSLSDHDIYYPNKIEHQVDIL 117

QY 534 LKEFLKDKTCLACVY-----TTNRNVPDGSILIANGYNWPEFSREKLTAMIAHFRMF 586
Db 118 NKLDNKDIIIGGYELIDDEKGNLSRYIKPDSVLPINKLN---ISLLPLLRGLIHGCSLLM 174
QY 587 TIRAWHLTDGFNENIENAVDYMFLKSEVGFKHLNKICYNRVLHGDNTSKLGIQK 646
Db 175 PAKYIYEVGIFNEALPTTQDYDLWFKIFRVAPIHFDESILIKSRFHSEGS-KKISNNHE 233
QY 647 N-----HFVVNQSLNRQGINNYNDFDLDDESKYIFNKTAQYQEMDKDLK 697
Db 234 ECVNLWSSFLHETBEEMIKMEGSPYL-----FTRTATFLSNNT-PYKKACDLANTWA 286
QY 698 LIQNKDAKIAVSIFYPNTLNGVLKLNNT-IEYNKNFIVILH----VDKNHLLTPDIKKE 752
Db 287 KOVLNDTKISVIIPVYNRINWAIEAKSVLIQTHKNFEIILDDGGSTDDISELTAICCKD 346
QY 753 -ILAYHKKHQVNNILLNNDISYTSNRLIKTEAHSNKLNSQLNCEYIIFDNHDSLFV 811
Db 347 KRIKYFHK-----KNEGPAAR-NLGKNAIGKYIAFLDSDDLFY 385
QY 812 KNDSYAYMKKYDVGWGNFSALTHDWIEKINAHPPFKLIKTYTFNDNLRSMNVKASQGMF 871
Db 386 K-DKIEIQKPFMEENN-IFSHTSYHKINE-----KGYIESVHSGLF 426
QY 872 MKYALPHELLTIKEVITSC 891
Db 427 SGNVFP-----QVIQTC 438
RESULT 3
AB1211
glycosyltransferases homolog lmo1090 [imported] - Listeria monocytogenes (strain EGD-
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AB1211
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloer
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehl
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1211
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-327 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAC99168.1; PID:g16410492; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo1090
Query Match 5.7%; Score 292.5; DB 2; Length 327;
Best Local Similarity 27.5%; Pred. No. 5.9e-09;
Matches 98; Conservative 44; Mismatches 133; Indels 81; Gaps 9;
QY 433 PLVSIYIPAYNCANYIQRCDVSALNQTVVDLEVCINDGSTDNLEVINKLGNNPRVRI 492
Db 3 PLVSIYIPAYNCANYIQRCDVSALNQTVVDLEVCINDGSTDNLEVINKLGNNPRVRI 60
QY 493 MSXPNGGIASASNAVSPAKYIYGQDSDDDYLEPAVELCFLKDKTCLACVYTTNRN 552
Db 61 FEKNGQATARNFGLDVATGDIYVMVYDDYISKNLVETCL-DTVQKTNADLVLFYSN 119
QY 553 VNPDS-----LIANGYNWPEFSREKLTAMIAHFRFTIRAWHLTDGF--- 597
Db 120 VNQEGKMQYIKRKGIKVLDAAGPTPNKFKYQADLWKG-----SRFPVGYEYDGLIPLV 173
QY 598 -----NENIENAVDY---DMFLKSEVGFKHLNKICYNRVLHGDNTSKLGIQKKN 647
Db 174 VTLKAKNPVKIQDALYIIITDRADSQSNIQOVDFLDVVI---MLENVETELKLGII---- 227

QY 648 HFVVVNSLRNQGINYNYDKFDLDSRKYIFNKTAETAEYQEEMDMLKDL---KLIONKDA 704
Db 228 -----YEEKDQALYLYIEHLYRLVL 249
QY 705 KTAVSIFYPNTLGLVKLNIIIEYKNKIFVILHVDKHNLPDIIKEILAFYHKH 760
Db 250 RKAIIYITNRQERKKLIKTIITQIEKFPNMGSPYQAGGKLTATLKKKALWLYLHH 305
RESULT 4
D95206
hypothetical protein Sp1771 [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: D95206
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfle,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: D95206
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-696 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK75845.1; PID:g14973268; GSPDB:GN00164; TIGR:SP4
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: Sp1771
Query Match 5.3%; Score 268.5; DB 2; Length 696;
Best Local Similarity 20.5%; Pred. No. 3.5e-07;
Matches 101; Conservative 96; Mismatches 181; Indels 115; Gaps 17;
QY 435 VSIYIPAYNCANYIQRCDVSALNQTVDLEVCICNDGSDTNTLEVINKLYGNPNRVRMS 494
Db 7 ITVIVPVYNNVLRKCLDSIIQTQYKNIIEVVVNDGSDASGEICKEFSEMDHRIILYIE 66
QY 495 KPNGGIASASNAVSAFAGYIIGDSDDYLEPDAVELCLK---ELFKDKTKLACVYTNR 551
Db 67 QENAGLSAARNTGLNNMNGNYVTFVDSDDWIEQDYVETLYKKIYEQADIAGVNGYSFNE 126
QY 552 N-----VNPDSLIANGYNWPFESREKLTMTAMIAHFRMFTIRAWHL-- 593
Db 127 SEGMYFHILGDSYEEKVDNVDSIFENLYE---TQEMKSFALISANGKLYKARLFEQLR 182
QY 594 -----TDGFNENIENAVDYDMFLKSEVGFKHLNKICY-NRVLHGD----- 634
Db 183 FOIGKLGEDGY-----LNQKVYL-LSE--KVLYLNKSLYAYRIRKGSLSRVWTEKWMH 232
QY 635 -----NTSIKGLGIOKKHNFVVVNSLRNQGINYNYDKFDLDSRKYIFNKTAE 695
Db 233 ALVDAMSERITLLANNGYFPLEKHLAYIRQMLEVSLAN---GOASGLSDTATY-----KE 283
QY 686 YOEMDMKDLKLIONKDAKIAVSIFYPNTLNLGLVKLNIIIEYKNKIFVILHVD-KNH 744
Db 284 FEMKQLLNQLSROESEKKAIVLAANYGYVDQVLTIKICVHNKSIREFYLHSDFPNE 343
QY 745 LTPDIKEILAF-----YHKHQNVLNNDISYNTNRLIKTEAHLNKNKLSQNLN 796
Db 344 WIKQLNKRLEKFPDSEINCIRVTSQIIS-CVKSDISYVTVFLRYFIADF---VQEDKALYL 398
QY 797 NCEYIIFNDHDSLFVKN-DSY-----AYMKKYDVGMN 827
Db 399 DCDLVVTKNLDDLFAFDLADYPLAAVYRDFGGRAYFQGEIFNAGVLLVNNFAFWKKENNTQK 458
QY 828 FSALTHTDWTIEKIN 840
Db 459 LIDVTNEHDKVD 471
RESULT 5

AD2189
hypothetical protein alr3067 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Notes: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AD2189
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iri,
R.; Kaneko, T.; Shimpō, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabat
Nakazaki, N.; Shimpō, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabat
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AD2189
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-323 <KUR>
A:Cross-references: GB:BA000019; PIDN:BAH74766.1; PID:g17132161; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr3067
Query Match 5.2%; Score 264.5; DB 2; Length 323;
Best Local Similarity 25.4%; Pred. No. 2e-07;
Matches 82; Conservative 64; Mismatches 116; Indels 61; Gaps 12;
QY 432 IPLVSIYIPAYNCANYIQRCDVSALNQTVDLEVCICNDGSDTNTLEVINKLYGNPNRVR 491
Db 10 VPIISVIPIPYNGEKTIIETIASVQHGTFLDIEIIVINDGSDTNTFELVRNIQDN--RLK 67
QY 492 IMSPKNGGIASASNAVSAFAGYIIGDSDDYLEPDAVEL---CLKEFLKDKTKLACVYT 548
Db 68 IFSYENGGLPVARNRGITHAVGQFIADIDADDLWTTKLELQFAALQEQY-PEAGLAYSWT 126
QY 549 TNRNVNPDGS-----LIANGYNWPFESREKLTMTAMIAHFRMFTI 588
Db 127 YKFAEADSYADESNFSAGDYAEELIKNLFQNGSN-PLIRAAIDSVGL----- 176
QY 589 RAWHLTDGFNENIENAVDYDMFLKSEVGFKHLNK--ICVNRVLHGDNTSIKKLGIOKK 646
Db 177 -----FDPTLKSCEDWDFYLRLEAKWQFALYKKAQIYRQ---SPTAMTSKLDYMEK 225
QY 647 NHFVVVNSLRNQGINYNYDKFDLDSRKYIFNKTAE--YQEMDMKDLKLIONKDA 704
Db 226 YSSIVIERAFNAAPPOQLHLKK-----QSLAWYKFTAQOCLKYNSHKLADIKLAA-KRL 279
QY 705 KIAVSIFYPNTL---NLGVKKL 723
Db 280 KMAITLYPKNLLLEDYTHGLIRKL 302
RESULT 6
A84114
glycosyltransferase BH3713 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: A84114
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masul, N.; Fujii, F.;
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: A84114
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-303 <STO>
A:Cross-references: GB:AP001519; GB:BA000004; MID:g10176109; PIDN:BAH07432.1; GSPDB
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH3713
Query Match 5.1%; Score 261.5; DB 2; Length 303;
Best Local Similarity 26.5%; Pred. No. 2.7e-07;
Matches 90; Conservative 60; Mismatches 95; Indels 95; Gaps 15;
QY 435 VSIYIPAYNCANYIQRCDVSALNQTVDLEVCICNDGSDTNTLEVINKLYGNPNRVRMS 494

Db 3 VSIILPHTNRAKLLKRALESTLNTQYKNIIEVIVVSDGSTNT-DIVMDYKRDSDRVNFIS 61
Qy 495 -KPNGGIASANAASVFAKGYIIGOLDSDDDLEPDVAVELCKEFLKDKTKLACVYT----- 548
Db 62 YHPAKGGYARNTGKNAKGFIAFLDSDDDMPDKLELQIKFNONANVLGVTTGVEII 121
Qy 549 ---TNRNVNPDGSLIANGYNWPE---FSREKL-----TTAMTAHHPFMTTIRAWHLTD 595
Db 122 YNFNKRNI-----KYISLPKKTGNLSKEILVANCIGTTSV-----MVRKNLITE 166
Qy 596 --GFENIENAVDVMFLKSEVGFKHLNKCICYNRVLHGDNTSIKKIGQKKNHFVVVN 653
Db 167 CGMPEDEKLARQDYDLWI-----RVC-----QKTLVGGVN 196
Qy 654 QSLNRQGINNYNDKFDL-DESKY-----IFNKTAE-----YQEMDMKL 694
Db 197 KPL-----VRYNYTTNKQISDDIKYSIAEYIDNKYVDLYSKVSEETRRKRRHSMTMLI 252
Qy 695 DLKIQNKDAKIA-----VSIFYPNTLNGVLKKNLNIIEY 729
Db 253 VNKALRNQSPKVARAYLAKNSFLKRTLITAILMYLSFLKY 292
RESULT 7
B97168
glycosyltransferase [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: B97168
R: Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: B97168
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333 <KUR>
A:Cross-references: GB:AB001437; PIDN:AAK80133.1; PID:g15025169; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC2175

Query Match 5.1%; Score 258.5; DB 2; Length 333;
Best Local Similarity 26.4%; Pred. No. 4.4e-07;
Matches 81; Conservative 54; Mismatches 125; Indels 47; Gaps 11;

Qy 434 LVSIIYPAYNCANYIQRCDVSALNQTIVVLEVCICNDGSDTNTLEVINKLYGNPNRVRIM 493
Db 1 MISVIMPVYNCBKYLESTESILKQTYRDFEIIVNDGSDNKSIDIINKYANDDNRIYVV 60
Qy 494 SKPNG-GIASANAASVFAKGYIIGOLDSDDDLEPDVAVELCKEFLKDK---TLACVYTT 549
Db 61 SRDNNMGVYSLNEGIDRAGSVARMADDDIALPERFERQIEYLNKKNKDVILACKVFA 120
Qy 550 NRVNPDGSL-IANGYNWPEFSREK-----LATAMTAHHPFMTTIRAWHLTDGTFNENIEN 603
Db 121 FGDVSRQKLEHWHYVNDLNNSESTESLFLENCYIAHPVSMKMSVLKALGGYNLYKR 180
Qy 604 AVDYDMFLKSEVG-KFKHLNKCICYNRVLHGDNTSIKKIGQKKNHFVVVNSLNROGIN 662
Db 181 TEDYNLWLRAIAGKYIAMLEKLMKIRLHNS-----KIHDA-- 219
Qy 663 YYNDRFDDLDDESKYIFNKTAEQEOMDKLKI---QNKDAKIA---VSIFYPNT- 715
Db 220 -EGFSSIRDIQSR-----LEYVKEKLKDKDFSVINGASNGGKIAYEKIKEVFPNAK 271
Qy 716 LNLVKK 722
Db 272 LNYIDK 278

RESULT 8
AG2188

hypothetical protein alr3062 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AG2188
R: Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigui
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG2188
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-321 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA874761.1; PID:g17132156; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr3062

Query Match 5.0%; Score 256; DB 2; Length 321;
Best Local Similarity 27.5%; Pred. No. 5.8e-07;
Matches 78; Conservative 50; Mismatches 118; Indels 38; Gaps 10;

Qy 432 IPLVSIYPAYNCANYIQRCDVSALNQTIVVLEVCICNDGSDTNTLEVINKLYGNPNRVR 491
Db 1 MPKVSWVIYPAYNAMPYLPETLESVLRQTYVHDFEVVVVNDGSSDNTTEWVSQL--DPLRK 58
Qy 492 IMSKPNGGIASANAASVFAKGYIIGOLDSDDDLEPDVAVELCKEFLKDKTKLACVYTTNR 551
Db 59 LISQANQGLAGARNTGIVNASGEYIAFDADDIWEFTKLAQVSLDENPTVGLVYTWVA 118
Qy 552 NVNPDGSLI-----ANGYNWPEFSREKLTAMTAHHPFMTTIRAWHLTDGTFNENIEN 604
Db 119 YDEQKSTGKTFKNOQEGYVWPQLTEHNVCEGVALVRRVCFEKMGL---FDRNLGSY 175
Qy 605 V-DYDMFLKSEVGKFKHLNK-ICYNRVLHGDNTSIK-----KLGIQKKNHFVVVNS 655
Db 176 VEDWDMWLRIATSYDFKVVKEALVYR--ORSNSASKNWEAMAHSAFVIERAFATASQD 233
Qy 656 L-----NROGIN-----YNDKFDLDDESKYIFNKTAIEY 686
Db 234 LQVLNKSXGFTYLCIAWKPLQSFQK--DYQKSREFCQQAIVY 275

RESULT 9
AB2190

hypothetical protein alr3073 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AB2190
R: Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigui
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AB2190
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-324 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA874772.1; PID:g17132167; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr3073

Query Match 5.0%; Score 254.5; DB 2; Length 324;
Best Local Similarity 26.2%; Pred. No. 7.1e-07;
Matches 101; Conservative 63; Mismatches 127; Indels 95; Gaps 16;

Qy 432 IPLVSIYPAYNCANYIQRCDVSALNQTIVVLEVCICNDGSDTNTLEVINKLYGNPNRVR 491
Db 1 MPKVSWVIYPAYNAMPYLPETLESVLRQTYVHDFEVVVVNDGSSDNTTEWVSQL--DPLRK 58

Db 1 MPKISVIIIPAYNAERTILETINSVLNQTFSDLEIIIVINDSGDRTVEVLQNV--DDARLK 58

QY 492 IMSKPNGGSIASNAASVAFKAYGIGQLSDDDYLEPDAVELCLKEFLDKTKTLACVYTTNR 551

Db 59 VYSYENSASGARNHGISHAVGDFISFDADDLWTPDKLEQLQSALNNHPEAGVAYSWTY 118

QY 552 NYNPDGSLIANGYNNPFSRREKLTMTAMIAHHFRMTIRAWHLTDGFNENIENAVDYDMEL 611

Db 119 TIDDKGELLKPE--PLYEGNYTDLLAN-----FLTNGSNPLIRKAA-----160

QY 612 KLSEVGKFKHLNKCYNRVLHGDNTSIIK-----LGIOKKNHFVVVNQSLNRQGINYY 664

Db 161 -IASICEF-----DITLRSGEDWYWLRLAYKWPVVVKQ-----HQILY- 199

QY 665 NYDKFDLDESRYIFNKTAEOEEMDKLKL-IQNKDAIYVIFYPNTLGLVK-K 722

Db 200 -----RRSVTSKSFILQ-----IIREASLAILDKAMKV-----LPLEQYLKHS 239

QY 723 LNNIIYKNKIFVIIHVDKNHLPDIIK-----KEILAFYHKKHQNIL 765

Db 240 LSNIRYNNVELYDSIN---NNSTVDIKYVIGNLLSYIRSRQTLKEI--YYKLIKIL 294

QY 766 LNDISYTSNRL-----IKTEAHLN 788

Db 295 LVIVLSPKLMSRLLOFIKSKOMKML 320

RESULT 10

B75096

glycosyl transferase PAB0772 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C:Accession: B75096

R:anonymous, Genoscope

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru

A:Reference number: A75001

A:Accession: B75096

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-298 <K>

A:Cross-references: GB:AJ248286; GB:AL096836; NID:g5458366; PIDN:CAB50071.1; PID:g545858

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB0772

C:Superfamily: Neisseria meningitidis glycosyl transferase A

Query Match 4.9%; Score 251; DB 2; Length 298;

Best Local Similarity 24.3%; Pred. No. 9.8e-07;

Matches 81; Conservative 66; Mismatches 136; Indels 50; Gaps 8;

QY 433 PLVSIYIPAYNCANYIQRVDSALNQTVDVLEVCICNDGSTNTLEVINKLYGNPNRVR- 491

Db 4 PIVSVIIPYNNRANLLRAIAISVFNOKFDFELIVVDDASTNTPEVSEI--EDGRIRY 61

QY 492 IMSKPNGGSIASNAASVAFKAYGIGQLSDDDYLEPDAVELCLKEFLK-DKTLACVYTTN 550

Db 62 IRLKNSGGPIARNIGIKKAKGRFIALDDDEWLPHRLEVQVRKFNELGKEGVYGGF 121

QY 551 RNVPDGLSIANGYNNPFSRREKLTMTAMIAHF-----RMFTIRAWHLTDGFNENIENAV 605

Db 122 YVVSQDGRIL--GKRLPK-HRGDIYSHLLKENFIGSPTLLIRRECFKAGLFPRLSSQ 178

QY 606 DYDMFLKLSEVGKFKHLNKCYNRVLHGDNTSIKKLGIOKKNHFVVVNQSLNRQGINYYN 665

Db 179 DWDMMIARIYKFDYVDIIIAKYVHGQISF----- 211

QY 666 YDKFDLDESRYIFNKTAEOEEMDKLKLQNKDAIYVIFYPNTLGLVKLN- 724

Db 212 -----NMKYIIPGRBLIRKHLDIWKNPKILSIHLSQMLLLLSNNTGKGLKLYIY 263

QY 725 --NIIYKNKIFVIIHVDKNHLPDIIKKEILA 755

Db 264 STAIAPLNLENYMIILLKALDSRTVEYIKRILS 296

RESULT 11

AH2026

hypothetical protein all1766 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C:Accession: AH2026

R:Kanakko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irlg Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabat DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AH2026

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-333 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA073465.1; PID:g17130856; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all1766

Query Match 4.7%; Score 241; DB 2; Length 333;

Best Local Similarity 26.4%; Pred. No. 4.1e-06;

Matches 70; Conservative 53; Mismatches 94; Indels 48; Gaps 7;

QY 432 IPIVSIYIPAYNCANYIQRVDSALNQTVDVLEVCICNDGSTNTLEVINKLYGNPNRVR 491

Db 8 LPIISVIIIPAYNCEKTIKTKIDSVLDQSFDTDFELIVINDGSQATLDIVSQI--EDSRIK 65

QY 492 IMSKPNGGSIASNAASVAFKAYGIGQLSDDDYLEPDAVELCLKEFLDKTKTLACVYTTNR 551

Db 66 IFSFENAGGVNSRNLNLAAGKFIPLDADDIWTNPKLESQLEALNNHGFHVAISWTD 125

QY 552 NVNPDGSLIANG-----YNNPFSRREKLT--TTAMIAHHFRMTIRAWHLT 594

Db 126 YIDEDGNFLISGRVTLNGDVYKFLFNFLNGSNPLICKEALIA-----L 172

QY 595 DGFNENIENAVDYDMFLKLSEVGKFKHLNKCYNRVLH--GDNTSIRKLGIOKKNHFVV 652

Db 173 GGFDESILKAAQDWMRLANKYSFV---AVPYVQILYRVSSNSLSNLVRQEKACKQVL 229

QY 653 -----NOSLNRQGINYYN 666

Db 230 EKAYQARPAIGNHILSLIANLYK 254

RESULT 12

T44330

glycosyl transferase homolog [imported] - Vibrio cholerae

C:Species: Vibrio cholerae

C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000

C:Accession: T44330

R:Yamasaki, S.; Shmizu, T.; Hoshino, K.; Ho, S.T.; Shimada, T.; Nair, G.B.; Takeda, Gene 237, 321-332, 1999

A:Title: The genes responsible for O-antigen synthesis of Vibrio cholerae O139 are c

A:Reference number: 222749; MUID:99453293; PMID:10521656

A:Accession: T44330

A:Status: preliminary;

A:Molecule type: DNA

A:Residues: 1-340 <YAM>

A:Cross-references: EMBL:AB012957; NID:g4115688; PIDN:BA033634.1; PID:g3721684

A:Experimental source: strain O22

C:Genetics:

A:Note: wblC

C:Superfamily: Neisseria meningitidis glycosyl transferase A

Query Match 4.7%; Score 239; DB 2; Length 340;

Best Local Similarity 25.4%; Pred. No. 5.4e-06;

Matches 78; Conservative 56; Mismatches 113; Indels 60; Gaps 11;

RESULT 1A

C;Genetics:

THIS PAGE BLANK (USPTO)

QY	20	LKLPKSAETGYGRKIVFQIIKEKELUSTNSYVSDEKKNVCDSSLDIATOLLNWKKL	79
		: : : : : : : : : : : : : : : : : : : : :	
Db	230	VEYFALAGSEYVRDATQGVIK-----RLELAETIETTERKAET-----VLVQKQLQEARR	280
		: : : : : : : : : : : : : : : : : : :	
QY	80	TLSESEK-----NSLKNNKWSITGKKSSNAEIRKVELVPKDPKDLVLAPLDHVNDET--	133
		: : : : : : : : : : : : : : : : : :	
Db	281	PLQLKRKLKFLNMLRAAKASPLPSRTAEFRFRSSAAKRD-----PMRDDLTQTLSGQ	332
		: : : : : : : : : : : : : : : : : : :	
QY	134	-----WYKNRKKSIG-----TKPVKNKIGLSIIPTFNRSRILDTITLACLVNQK	177
		: : : : : : : : : : : : : : : : : :	
Db	333	GFWTYEAIVRGWGQRQAGLAGRSELVVRNQLONGPLISVVVVVPPNPDPAIIEMIESVRAQ	392
		: : : : : : : : : : : : : : : : : : :	
QY	178	TNYPFEVVVADDSKSE--NLLTIVOKY--EOKLDIKYVRKQDYGQLCAVRNLGLRFTAKYDF	235


```
DR EMBL; AF031959; AAC32401.1; ALT_INIT.
DR PIR; B35391; B35391.
DR InterPro; IPR001173; Glycos_transf_2.
DR Pfam; PF00535; Glycos_transf_2; 1.
KW Hypothetical protein; transferase; Glycosyltransferase;
KW Complete proteome.
SQ SEQUENCE 322 AA; 36388 MW; 753C2FB59327D968 CRC64;

Query Match          3.9%; Score 199.5; DB 1; Length 322;
Best Local Similarity 26.4%; Pred. No. 0.00034;
Matches 81; Conservative 48; Mismatches 101; Indels 77; Gaps 16;

QY 435 VSIYIPAYNCANYIQRCDVSALNQTVDVLEVCICNDGSDNTFVINKLYGNP-RVRIM 493
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 ISVIISNYARYLSRAINSVLAQTHSDIEIVIVDDGSDNSRDVITQLQEQAPDKIKPI 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 494 SKPNGGIASANAASFAKGYIIGOLDSDDYLEPDAVELCLKEFL----- 538
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 FOANOGGGAFAAGFAAATGEVAFDADVDVWPKHKLQRIVEFQTSDDVGYMHHLDIID 122
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 539 -KDKTLACVYTNRNVPD-GSLI---ANGYNWP-----EFSREKLTAMIAHHPFMTI 588
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 GNDKTIQDASTGPKLSEDLASVILQTGNACFPPTSGLAYREVLE-----KVFP 174
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 589 -----RAWHLTDGFNENIENAVDY-DMFLKLSEVGKFKHLNK-ICYNRVLHGDNTSI--- 638
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 175 DPVKRWI--ADG-----CIYCTAFL-----GKIKTLQENLAYYRI-HGANNHMSAA 219
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 639 -----KKLGIOKKNHFFVVVNSLNROGINYNYDKFDLDSRKYIFNKTAHYQEE 689
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 220 SATSQEAKSQAGIEMTQY--INDFLVRIGY-----ARVDLSRNLYQRRTRYQORS 270
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 690 MDMLKDL 696
   : : : :
Db 271 QWDLREV 277
```

```
DR Pfam; PF00535; Glycos_transf_2; 2.
KW Lipopolysaccharide biosynthesis.
SQ SEQUENCE 1275 AA; 139596 MW; 3AF9662A10A140F1 CRC64;

Query Match          3.9%; Score 197.5; DB 1; Length 1275;
Best Local Similarity 20.9%; Pred. No. 0.0026;
Matches 84; Conservative 71; Mismatches 176; Indels 71; Gaps 14;

QY 427 SHHRIPLYSIVIPAYNCAN-YIQRCDVSALNQTVDVLEVCICNDGSDT-DNTLEVLNKL 484
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 313 ARLSRPLISLVTVPVHDASEAFURECLASVSSQVYADWEWLLVDDASTAPHLARIIEAA 372
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 485 GNNPRVRIM-SKPNGGIASANAASFAKGYIIGOLDSDDYLEPDAVELCLKEFLDKRTL 543
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 373 ERESRIVLTASSEGDTARATNEGFAACRGDFVGFELGAEDTLSPHALAEVALAFLAQPEL 432
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 544 ACVYTNRNVPDGSGLIANGYNWPEFSREKLTAMTAHIFRMTITRAWHLTDCGFENIEN 603
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 433 ALLYTDDEDGLDAQGHRSAPEFK-PDWSPDLLRSVDVVRHFLVVRRETQAQVGLREGFDG 491
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 604 AVDYDMFLKLE-VGKFKHLNKICYNRVLHGDNSTIKK---LGIOKKNHFFVVVNSLNRO 659
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 492 AGHDLMKLSEATSSIGHITEPLYH-AREGSAASASRGAGLDTATKAGVRALSEHLARQ 550
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 660 GINYNYDKFDLDSRKYIFNKTAHYQEEEMDKLKLQNKDKAKIAVSIFYPNTLNLGL 719
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 551 G-----ESAEV-----TSPAIQYRVYP----- 569
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 720 VKKLNIIYNNKIFVILHVDKNHLYTPDIKKEILA--FYHKHQNVLNND----- 769
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 570 -----VRGTPKVSIIIVPKDRPDLRLTLVDSLLAQTRYPHFVLLVSNNSRTPETFA 622
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 770 ISYVTSNRLIK-TEAHLN---INKLSQLNLCNEYIFDNHD 807
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 623 LEQWDPRLVKLTWDHPFNYPAINNAKQAQSGELLFLNND 664
```

```
RESULT 8
Y025_MYCPN
ID Y025_MYCPN STANDARD; PRT; 299 AA.
AC P75086;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative glycosyl transferase MG025 homolog (EC 2.-.-.-)
DE (B01_orf295V)
GN MPN028 OR MPl26.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC STRONG, TO M.GENITALIUM MG025.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000015; AAB95774.1; -.
CC InterPro; IPR001173; Glycos_transf_2.
DR Pfam; PF00535; Glycos_transf_2; 1.
```

```
KW Hypothetical protein; Transferase; Glycosyltransferase;
KW Complete proteome.
SQ SEQUENCE 299 AA; 34935 MW; 9A3AF2F099283565 CRC64;

Query Match          3.8%; Score 193.5; DB 1; Length 299;
Best Local Similarity 24.5%; Pred. No. 0.00066;
Matches 74; Conservative 38; Mismatches 127; Indels 63; Gaps 10;

QY 434 LVSIYIPAYNCANYIQCVDLSAL--NOTVVDLEVCINCDSSTNTLEVINKLYGNPRVR 491
DB 1:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 492 IMSPNGGIASASNAV--SFAKYYIGOLDSDDDYLEDPAVELGCKFLKDTLACVVT 549
DB 6 LFTVLIIPYNGQVIPRALDLSLLGGEYFTKTLIVLNDGSTNDTKTQVPEYTOQYSN 65
QY 492 IMSPNGGIASASNAV--SFAKYYIGOLDSDDDYLEDPAVELGCKFLKDTLACVVT 549
DB 66 YLEKPNMGWGVNPFVQKONLAKQYITVLDSDDFLANAFQVAAHFGHDMIVSAFVC- 124
QY 550 NRVNPDGSLIANGYNWPEFSREKLTAMIAH---FRMTIRAWHLTDGFNENIENAVD 606
DB 125 --YISPKRRRLKPYEGKTGVIEQTKLRTPHSQPLAKFYRHEIFHLDLPLKEL---FY 179
QY 607 YDMFKLSEVGKFKHLNKC-----YNRVLHGDNTSIKGLGKQKHHFVVVYNQSLNRQGI 661
DB 180 QDCLLHNAINKVQSVFVCEPLAVYATRFQGNSTMP----- 217
QY 662 NYNYTKFDDLDESRYIFNKTAEYQOEEDMLKDLKLQNKDAKIAVSIFYPNTLGLVK 721
DB 218 -WNNNAKFE-----QAWCDLLKQNNLY---GAGIVIIY--TMLPGFLK 254
QY 722 KL 723
DB 255 EL 256

RESULT 9
GGAA_BACSU STANDARD; PRT; 446 AA.
AC P46917;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Minor teichoic acids biosynthesis protein ggaA.
GN GGAA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC Freymond P., Karamata D.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriell S., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerion I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,
RA Ghim S.-Y., Glaser P., Goffeau A., Golligly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaert-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maestl D., Nakai S., Noback M.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Oudega B., Park S.H.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara S., Prescott A.M.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Rapoport G.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
```


RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RL burgdorferi".
CC Nature 390:580-586(1997).
CC -!- FUNCTION: REQUIRED FOR EFFICIENT DNA REPAIR; IT CATALYZES THE
CC UNWINDING OF DOUBLE-STRANDED DNA AND THE CLEAVAGE OF SINGLE-
CC STRANDED DNA AND IT STIMULATES LOCAL GENETIC RECOMBINATION.
CC ALL OF THESE ACTIVITIES REQUIRE CONCOMITANT HYDROLYSIS OF ATP
CC (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage (in the presence of
CC ATP) in either 5'- to 3'- or 3'- to 5'-direction to yield 5'-
CC phosphooligonucleotides
CC -!- SUBUNIT: CONSIST OF THREE SUBUNITS; RECB, RECC AND RECD
CC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE HELICASE FAMILY. UVRD SUBFAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE001164; AAC66981.1; .
CC HSSP; P56255; 1PJR.
CC TIGR; B06233;
CC InterPro; IPR004586; RecB.
CC InterPro; IPR000212; UvrD-helicase.
CC Pfam; PF00580; UvrD-helicase; 1.
CC TIGRPFAMS; TIGR00609; recB; 1.
CC Hydrolase; Nuclease; Exonuclease; Helicase; ATP-binding;
CC DNA repair; Complete proteome.
CC NP_BIND 18 25 ATP (POTENTIAL).
CC SEQUENCE 1169 AA; 137828 MW; B61D63C1C959B91F CRC64;
Query Match 3.6%; Score 183; DB 1; Length 1169;
Best Local Similarity 19.3%; Pred. No. 0.014;
Matches 175; Conservative 138; Mismatches 321; Indels 272; Gaps 44;
QY 35 VEFQIIKCKEKLSTNSVYSEDKNSVCDDSLDIATQLLLSNVKKLTLSSESK--NSLAKN 92
DB 313 VEYKTLK-----YTELKTKTI-KSTNTIDQNYIISNLNKLKSEDKLLNAIKNR 362
QY 93 WKSITGKKSNAEIRKVEL--VPKDFPKDLVLAPLDHV-----NDFTWY----KNR-- 138
DB 363 YKILIDPAQLSLQIEFLFKLTAGIKLFIADPKQIYSEFRKADISFYNKEIKKNIN 422
QY 139 -----KKSIGIKPVNK---NIGLSIIPTFNRSRLDIYLAIVNOKTNP 181
DB 423 TDARIVLKNHRSKKLIG--PLNKFIFNNYNAIADETEK---IDFTNS-LPNOKNKN 476
QY 182 FEWVYADGSKENLTIYQVYEQ-----KLDIKVVRQKDYGYQLCAVRNGLRITAKYDF 235
DB 477 KIVNGQEIEGINIITTFNESEDIYQKALTIKYLLA--YCK---IAENNKIRIKMQD 531
QY 236 VSILDCDAPQOLWVHSLYTELLEDNDIVLIG---PRKYVTHNTITAEQFLNDP----- 286
DB 532 IKVL-----CRGKNEINLIDKALKKEQIQIOT-NKTQEKFLTKTFSEIF 573
QY 287 YLIESLPETATNNPS-ITSGNISIDWRLHEFK-KTDNLRCDSPFRFVA----- 336
DB 574 YIIKCLDRKOSPKTLNVIUSSKILNVPWNLQRILLKQKICLIEFIENIIVLLEKNEIT 633
QY 337 -----GNVAFSEWLNKVGWF--DEEFNHWGGEDVEFYGLRFAKGFVRVIDGMAIHOE 389
DB 634 LINAINKITFEKNLWIKIANITKQKIIIEWAKNKINY-----KGLLI----- 675
QY 390 PPKENETREACKSITLKIVKEKVPYIYKLLPIEDSHHRIPLVSIPIPAYNCANYIQ 449
DB 676 ---KEGKLENLKYETTLTLEISK-----ITHKEQNIQSLS 708

QY	450	RCVDSALNQTVDVLEVCINCDSGTNTLEVI-----NKLIGNPVRIRMSKP--NGGIASA	503
Db	709	TLESIIINEPEIEEKININNDNESTELMTIHKSKGLGMNIVELLNTTPIENSFFSK	768
QY	504	SNAASFAGYIYGQLDSDDY-LEPDAVELCLKEFLKDKTKLACVYTT-----	549
Db	769	KN---QFYFYQDGKTEYDFKLEENKKYARLKILSEKNIFYVGATRAKFALEFIKINS	825
QY	550	-----NRNVN-----PDGSLIANGYNWPE	568
Db	826	ITSKLELEIAKFTIDDIKHDFNIHFEGKRFNKKYNTNVNTKLPPKPIKNMFK-KE	884
QY	569	PSREKLTAMIAHFRMFITRAHWLTDGNGENIENAVDYDMFL-----KLSEVGKFKH--	621
Db	885	YTSFSSLTAAQHAKFEFYENYDF---KNINYEKETEDELPGLEETLPKGRDIGNILHAA	941
QY	622	LNKICYNRV-----LHGDNSTSKLIGIOKKN---HFVVNOSLNRRQNGINYNYD-----	667
Db	942	MEELIFSTAKDTFDFNKKKNWIEIEKIQOKINSNLNTIEIQNSLAKMIYNILTYNRAIN	1001
QY	668	---KFDDLDESKYIFPNKTAEQEEMDMKLKLLKLNK---DAKTAVSIFYPNTLNGLV	720
Db	1002	TRLCDIEELQKEMEFLEIKINPEFOKQ-----KYLFDRKHFDLHKLKSGYLKGIYDLI	1054
QY	721	KKLAN---ITEYNKNIEFIIHLHYDKN-----HLTPDIIKEELAPYHKKHV-----	766
Db	1055	FKANKIYILDYKTN-----YLGKKNKEDYNITNLENTIKKEYDYLQYKIYALGIKKILF	1108
QY	767	NDISY 772	
Db	1109	KNKREY 1114	
RESULT 12			
Y087_BUCAI		STANDARD;	970 AA.
ID	Y087_BUCAI		
AC	P57189;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Hypothetical protein BU087.		
GN	BU087.		
OS	Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum		
OS	symbiotic bacterium).		
OC	Bacteria; Proteobacteria; gamma subdivision; Buchnera.		
OX	NCBI_TaxID=118099;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Tokyo 1998;		
RX	MEDLINE=20445173; PubMed=10993077;		
RA	Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;		
RT	'Genome sequence of the endocellular bacterial symbiont of aphids		
RT	Buchnera sp. APS.';		
RL	Nature 407:81-86(2000).		
CC	-1- SUBUNIT: SOME, TO E.COLI YTFN.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announcement/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; AP001118; BAB12807.1;		
KW	Hypothetical protein; Transmembrane; Complete proteome.		
FT	TRANSMEM 12 32		
FT	POTENTIAL.		
SEQ	SEQUENCE 970 AA; 114477 MW; 16B7BADB129F422E CRC64;		
Query Match 3.5%; Score 177.5; DB 1; Length 970;			
Best Local Similarity 20.3%; Pred No. 0.022;			
Matches 202; Conservative 134; Mismatches 302; Indels 357; Gaps 57;			

Query Match 3.5%; Score 177.5; DB 1; Length 970;
Best Local Similarity 20.3%; Pred. NO. 0.022;
Matches 202; Conservative 134; Mismatches 302; Indels 357; Gaps 57;

```
QY 67 IATQLLSNVKKLTSESEKSNLKNWKSITGKKSNAEIRKVELVPKDPKDLVLAPL 126
Db 126 IETKNVILSLKKNATSNFSLPS-----KISKNIFFIKYPIILKKIHADKIL--- 144
QY 127 DHVNDFTYKRNKSL-----GIKPVKNIGLSIIPTNRSRILDTITLACLVNQKNTNPF 182
Db 182 FTSPKRVSVSLDLSGKILRNNI---IFSPTY-----INTIHVSAAKFNF-- 187
QY 183 EVVADDSKENLL---TIVQKYEOKLDIKVVRKDY-----GYQLCAVRNGLRLTA 231
Db 231 EKNILNKSTIIKNFNK---IKKIYSFYSSTNOTKKNFPL-----NIYLKSL 232
QY 232 K-----YDFVSILDCDMPAQLVHVSHTLELEDNDIVLIGPRKYVDTH-----NITA 279
Db 279 KCKNTQFIDYEVKNLQVEL-----QANIENNILQINKMK-VDSFLKMNISVG 279
QY 280 EQLNDPILIESLPAT-----NNPFSITSKGNISLDWLEHFKKTDN----- 323
Db 323 KVFENNDYSISCVMSKTVIPSLYNKSNFQKANKFNVDHQL-IFKLSKDLNKMKNGL 338
QY 334 --LRLCDSPFRFV--AGNVA--FSKWLKNKVGWDFDEEFNHGWDVEFGYRLFAGKGF 376
Db 376 VLFNFSDFPF--FIKLRNSLSCVTKKNYIFKLKSF-----GVLKGI-----INNY 383
QY 377 FRVIDGMAIHQEP-----GKENETEREAGKSITLKIYKVPYIYKLLPIED----- 426
Db 426 FFSKNIPTLODLPPIFIDIOGR-GDLNNIFLKKINFRPIKQK--KFKVKVHPEDIKY 440
QY 427 -----SHIHRIPLSYIYPAYNCANYIQRVDSALNQTVDLEVC 466
Db 466 NOYILKLGQINITGKSDRHTHYVHPIDLY-----ANIMKK----- 478
QY 467 ICNDGSTDLTLEVINKLGN-----PRVRIMSKPN-----GGIASASNAVSPAKGY 515
Db 515 KLSILGALYKKNFNFTETGGINLLGKNKLYLRSGKKNYIYSSI---Y 525
QY 516 IQLSDSDYLEPDAVELCLKEFLDKTLACVYTTNRNVP--DGLSIANGYNWPE--FSR 571
Db 576 ANNL--DYFFPK-----LOGRMAKVNFNGNRPFIISKILARDLNNNIFKN 573
QY 572 EKLTTA-----MTAH-----HFRMTIRAWHLTDTGFNENIENAVDYDMFLKSEVG 617
Db 574 IKVLGTGINNTFSKMLIYANKIHFKFYINTLHIQYTSNNHKQN---FSELLKSNRL- 629
QY 618 KFKHLNKI---CYNRVLHGDNTSIIKLGIO-----KKNHFVVVNNQSLNROGINNY 666
Db 630 ---HLNLINGAFNKTGHWHGFFKINIRTFWQGVQVTAKKNNFIHYDHSNITNFKYQKS 686
QY 667 DKFDDLDESRY-----IFNKT-AEYQEDMDLKDCLKIQNK----- 702
Db 687 IKKRNCFSSFLYNNKMSFNLNRSFISFESKLSINAKLILGKMSIDGAIPLKGNNTK 746
QY 703 -DAKLAVSIFYPN-----TLNGLVK-----KLNIIENYKNIF-----VILHVDK 742
Db 747 LEKKLKKIFIQIDFFKISMN-LIKNDPKSKWIIKKKKLNNKIFGLNIIDIYNNK 805
QY 743 NHLTDPDIKEILAEVAKOVNILLANDLSYTSN-----RLIKT----- 782
Db 806 N-----IKGEFI--FYKPPFP-----INFTTNEKESVGKFSQKIFGLTYQPKVLAD 853
QY 783 AHLNLI-----NKLSQLNLNCEYIIFDNDHSLFV-----KNDS----- 815
Db 854 VHFKNIFIRSNILAYITLFFPYFLGKVDNIKINOE-IMNKGNIITLTKLPFKKNSADI 912
QY 816 ---YAYMKKYDVMNFSALTHWDIEKINAHPPFKK 847
Db 913 ENNIAFNKSKISVLIFPKIKVFESKLNHLFLFSK 947

RESULT 13
Y025_MYCGE
ID Y025_MYCGE STANDARD; PRT; 298 AA.
AC P47271; Q49359;
```

```
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative glycosyl transferase MG025 (EC 2.4.1.10).
GN MG025.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37; PubMed=7569993;
RA MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bost K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium."
RL Science 270:397-403(1995).
RN [2]
RP SEQUENCE OF 172-298 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RA MEDLINE=94075230; PubMed=8253680;
RA Peterson S.N., Hu P.-C., Bost K.F., Hutchison C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
RT sequencing."
RL J. Bacteriol. 175:7918-7930(1993).
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC STRONG, TO M.PNEUMONIAE MPN028.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U36682; AAC1241.1; -
DR EMBL: U02253; AAC12517.1; -
DR TIGR: MG025; -
DR InterPro: IPR001173; Glycos_transf_2.
DR Pfam: PF00535; Glycos_transf_2; 1.
KW Hypothetical protein; Transferase; Glycosyltransferase;
KW Complete proteome.
FT CONFLICT 200 201 EP -> DA (IN REF. 2).
SQ SEQUENCE 298 AA; 35044 MW; B7BCCC73EB2CA023 CRC64;

Query Match 3.5%; Score 176.5; DB 1; Length 298;
Best Local Similarity 21.8%; Pred. No. 0.0055;
Matches 80; Conservative 40; Mismatches 118; Indels 129; Gaps 14;

QY 434 LVSIIYIPAYNCANYIQRVDSAL--NQTVVDLEVCINDGSTDTNTLEVINKLYGNPRVR 491
Db 6 LFTVIPTTNCQYIKKALDLSLLQNEFLKTLQVTLVNDGSLDNTKEVSDYLIKYSNIS 65
QY 492 IMSKPNNGGTASASNAV--SFAKGYIIGQLSDSDYLEPDAVELCLKEFLDKTLACVYT- 548
Db 66 YFEKTNGNGSVINYYKKNKALGOYITVLDSDDYFLKDSFKKVRFFGHDMITGAFYCY 125
QY 549 TNRNVNPDGSLIANGYNWPEFSREKLTATIAHHFRMTIRAWHLTDTGFNENIENAVDYD 608
Db 126 INEN-----KTRFLKPYFGK---TGVIKREHTKLRTPHS----- 155
QY 609 MFLKLSEVGKFKHLNKCYNRVLHGDNTSIKLGIOKKNHFVVVNNQSLNROGINNYNDK 668
Db 156 ----QPIAKF-YSNKLFYE--LH----- 171
QY 669 FDDLDESRYIIFNKTAEYQEDMDLKDCLKIQNKDAKIAVSIFY----- 712
Db 172 --DLKE--KLFFQDCLMYHDAINRVE-----SVFYREPLAVNWFSTRPGNS 213
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2003, 02:13:54 ; Search time 121 seconds
(without alignments)
1643.268 Million cell updates/sec

Title: US-09-842-484A-2
Perfect score: 5089
Sequence: 1 MNTLSQAIKAYNSNDYELAL.....SAKGENIPVKNFINSITL 965

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5
Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5089	100.0	965	2 Q9KJ99	Q9KJ99 pasteurilla
2	4802	94.4	965	16 Q9CMP0	Q9CMP0 pasteurilla
3	4791	94.1	965	2 Q9AHL6	Q9AHL6 pasteurilla
4	4493.5	88.3	972	2 Q85457	Q85457 pasteurilla
5	4486.5	88.2	972	2 Q68389	Q68389 pasteurilla
6	294	5.8	604	16 Q921F9	Q921F9 rickettsia
7	292.5	5.7	327	16 Q8Y828	Q8Y828 listeria mo
8	292	5.7	337	16 Q9CLR9	Q9CLR9 pasteurilla
9	287.5	5.6	350	2 Q9ALS8	Q9ALS8 campylobact
10	279.5	5.5	326	16 Q8XN37	Q8XN37 clostridium
11	276.5	5.4	706	2 Q88090	Q88090 enterococcu
12	268.5	5.3	696	16 Q97P72	Q97P72 streptococc
13	265.5	5.2	348	16 Q8XN34	Q8XN34 clostridium
14	264.5	5.2	323	16 Q8YSL7	Q8YSL7 anabaena sp
15	261.5	5.1	303	16 Q9K6L6	Q9K6L6 bacillus ha
16	259.5	5.1	323	2 Q86893	Q86893 streptococc

17	259	5.1	342	16 Q8XN57	Q8XN57 clostridium
18	258.5	5.1	333	16 Q97H38	Q97H38 clostridium
19	257	5.1	316	2 Q06035	Q06035 lactococcus
20	256	5.0	321	16 Q8YSM2	Q8YSM2 anabaena sp
21	256	5.0	330	16 Q8XN54	Q8XN54 clostridium
22	254.5	5.0	324	16 Q8YSL1	Q8YSL1 anabaena sp
23	254	5.0	315	2 Q93TI5	Q93TI5 streptococc
24	253	5.0	321	2 Q9AFH3	Q9AFH3 streptococc
25	251	4.9	298	17 Q9U2I6	Q9U2I6 pyrococcus
26	249.5	4.9	283	2 Q93TI7	Q93TI7 streptococc
27	247	4.9	324	2 Q9AQJ0	Q9AQJ0 streptococc
28	246	4.8	322	2 Q9RG46	Q9RG46 streptococc
29	241.5	4.7	281	2 Q9ZGK2	Q9ZGK2 leptospira
30	241	4.7	324	2 Q8VLB0	Q8VLB0 streptococc
31	241	4.7	323	2 Q87L82	Q87L82 streptococc
32	241	4.7	333	16 Q8YW50	Q8YW50 anabaena sp
33	241	4.7	582	2 Q9AEE2	Q9AEE2 leptospira
34	241	4.7	582	2 Q8VTX6	Q8VTX6 leptospira
35	239	4.7	281	2 Q9S4F8	Q9S4F8 leptospira
36	239	4.7	340	2 Q87159	Q87159 vibrio chol
37	238.5	4.7	323	2 Q85000	Q85000 streptococc
38	238.5	4.7	324	2 Q86046	Q86046 streptococc
39	238	4.7	306	2 Q07339	Q07339 streptococc
40	238	4.7	316	16 Q8YSL6	Q8YSL6 anabaena sp
41	237.5	4.7	257	16 Q9K6R6	Q9K6R6 bacillus ha
42	237.5	4.7	306	2 Q9RP62	Q9RP62 escherichia
43	237	4.7	344	16 P71059	P71059 bacillus su
44	236.5	4.6	577	2 Q9S4F7	Q9S4F7 leptospira
45	236.5	4.6	1238	12 Q9EMP3	Q9EMP3 ansacta moo

ALIGNMENTS

RESULT 1

ID	Q9KJ99	PRELIMINARY;	PRT;	965 AA.
AC	Q9KJ99;			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Chondroitin synthase CS.			
OS	Pasteurella multocida.			
OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;			
OC	Pasteurella.			
OX	NCBI_TaxID=747;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=P4679;			
RX	MEDLINE=20379058; PubMed=10818104;			
RA	DeAngelis P.L., Padgett-McCue A.J.;			
RT	"Identification and Molecular Cloning of a Chondroitin Synthase from			
RL	Pasteurella multocida Type F.;"			
RL	J. Biol. Chem. 275:24124-24129(2000).			
DR	EMBL; AF195517; AAF97500.1; -.			
DR	HSSP; P39621; 10GO.			
DR	InterPro; IPR001173; Glycos_transf_2.			
DR	Pfam; PF00535; Glycos_transf_2; 2.			
SQ	SEQUENCE 965 AA; 111392 MW; AFD55CD59D35C8C CRC64;			
Query Match	100.0%;	Score 5089;	DB 2;	Length 965;
Best Local Similarity	100.0%;	Pred. No. 8.1e-256;		
Matches 965;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1 MNTLSQAIKAYNSNDYELALKLFKSAETYGKRIKKEKLSSTNSYVSEDKNSV 60			
Db	1 MNTLSQAIKAYNSNDYELALKLFKSAETYGKRIKKEKLSSTNSYVSEDKNSV 60			
QY	61 CDSSLDIATQLLSNVKLTLSSEKNSLKNWKSITGKSENAEIRKVELVPKDFPKDL 120			
Db	61 CDSSLDIATQLLSNVKLTLSSEKNSLKNWKSITGKSENAEIRKVELVPKDFPKDL 120			
QY	121 VLAPLPDHVNDFTWYKNRKKSLGKIPVKNKNIGLSIITPFNRSRILDITLACLIVNQKTNV 180			

Db	121	VLAPLPHVNDFTWYKRRKSLGKIPVKNKIGLSIIPTFNRSLDITLACLNVQNTNY	180
Qy	181	PFEVVVADGSKENLLTIVQYEQKLDIKYVRQKDYGYQCAVRNGLRTAKYDFVSILD	240
Db	181	PFEVVVADGSKENLLTIVQYEQKLDIKYVRQKDYGYQCAVRNGLRTAKYDFVSILD	240
Qy	241	CDMAPOQLWHSYLTLELNDIVLIGPRKYVDTHNTAEQFLNDPVLIESLPETATNNN	300
Db	241	CDMAPOQLWHSYLTLELNDIVLIGPRKYVDTHNTAEQFLNDPVLIESLPETATNNN	300
Qy	301	PSITSKGNISLDWLEHFKKTDNLRCDSPFRYFVAGNVAFSKWLNKVGWDFEENHWG	360
Db	301	PSITSKGNISLDWLEHFKKTDNLRCDSPFRYFVAGNVAFSKWLNKVGWDFEENHWG	360
Qy	361	GEDVEFGYRLFPAKCFRVIDGGMALHQPPEPKENETEREAGKSITLKVKEKPYIYRK	420
Db	361	GEDVEFGYRLFPAKCFRVIDGGMALHQPPEPKENETEREAGKSITLKVKEKPYIYRK	420
Qy	421	LLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVDLEVCICNDGSTNTLEVI	480
Db	421	LLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVDLEVCICNDGSTNTLEVI	480
Qy	481	NKLYGNPRVRIMSKPNGGSIASASNAAVSFAGYIYGQDSDDDYLEPDAVELCLKEFLKD	540
Db	481	NKLYGNPRVRIMSKPNGGSIASASNAAVSFAGYIYGQDSDDDYLEPDAVELCLKEFLKD	540
Qy	541	KLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHHFRMFTIRAWHLTDGFNEN	600
Db	541	KLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHHFRMFTIRAWHLTDGFNEN	600
Qy	601	TENAVDYDMFLKLSVGEKFLHNLKICYNRVLHGDNTSIKLGIOKKHFFVNVQSLNRQG	660
Db	601	TENAVDYDMFLKLSVGEKFLHNLKICYNRVLHGDNTSIKLGIOKKHFFVNVQSLNRQG	660
Qy	661	INYNYDKFDDDESRRKYIFNKTAIEYQEEIDILKQIKVQRKAKVAISIFYPNTLGLV	720
Db	661	INYNYDKFDDDESRRKYIFNKTAIEYQEEIDILKQIKVQRKAKVAISIFYPNTLGLV	720
Qy	721	KLNNIIEYKNKIFVILLHVDKNHLPDIIKKEILAFYHKHQNILLNDISYTTSNRLIK	780
Db	721	KLNNIIEYKNKIFVILLHVDKNHLPDIIKKEILAFYHKHQNILLNDISYTTSNRLIK	780
Qy	781	TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAMKDYDGMNFSALPHDWEIKIN	840
Db	781	TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAMKDYDGMNFSALPHDWEIKIN	840
Qy	841	AHPPFKLLIKTYFNNDLRSMNVKASQGMFKYALPHELLIIEKITSCOSIDSVPYI	900
Db	841	AHPPFKLLIKTYFNNDLRSMNVKASQGMFKYALPHELLIIEKITSCOSIDSVPYI	900
Qy	901	NTEDIWFQFALLILEKKTGHVFNKSTLTLYMPWERKLOWTNEQISAKKGENIPVNFII	960
Db	901	NTEDIWFQFALLILEKKTGHVFNKSTLTLYMPWERKLOWTNEQISAKKGENIPVNFII	960
Qy	961	NSITL 965	
Db	961	NSITL 965	

RESULT 2

Q9CMP0 PRELIMINARY; PRT; 965 AA.

ID Q9CMP0; AC Q9CMP0; DT 01-JUN-2001 (TrEMBLrel. 17, Created); DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update); DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update); DE Hypothetical protein PM0775; GN .PM0775; OS Pasteurella multocida; OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; OC Pasteurella; NCBI_TaxID=747; OX

RP	SEQUENCE FROM N.A.	
RC	STRAIN=PM70;	
RX	MEDLINE=21145866; PubMed=11248100;	
RA	May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;	
RT	*Complete genomic sequence of Pasteurella multocida Pm70.*;	
RL	Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).	
DR	EMBL: AE006116; AAK02859.1; -	
DR	HSSP: P39621; LQGO.	
DR	InterPro: IPR001173; Glycos_transf_2.	
DR	Pfam: PF00535; Glycos_transf_2; 2.	
KW	Hypothetical protein; Complete proteome.	
SQ	SEQUENCE 965 AA; 111600 MW; 9C4B2CF80E1A6BD7 CRC64;	
	Query Match 94.4%; Score 4802; DB 16; Length 965;	
	Best Local Similarity 93.0%; Pred. No. 6e-241;	
	Matches 897; Conservative 42; Mismatches 26; Indels 0; Gaps 0;	
Qy	1	MNTLSQAIKAYNSNDYELALKLFKSAETYGKIVEQIICKKEKLSNSTSVSDKNSV 60
Db	1	MNTLSQAIKAYNSNDYELALKLFKSAETYGKIVEQIICKKEKLSNSTSVSDKNSV 60
Qy	61	CDSSLDATQLLSNVKKLTLSSEKNSLKNKWSITGKSENAEIRKVELVPRDKDL 120
Db	61	CDSSLDATQLLSNVKKLTLSSEKNSLKNKWSITGKSENAEIRKVELVPRDKDL 120
Qy	121	VLAPLPHVNDFTWYKRRKSLGKIPVKNKIGLSIIPTFNRSLDITLACLNVQNTNY 180
Db	121	VLAPLPHVNDFTWYKRRKSLGKIPVKNKIGLSIIPTFNRSLDITLACLNVQNTNY 180
Qy	181	PFEVVVADGSKENLLTIVQYEQKLDIKYVRQKDYGYQCAVRNGLRTAKYDFVSILD 240
Db	181	PFEVVVADGSKENLLTIVQYEQKLDIKYVRQKDYGYQCAVRNGLRTAKYDFVSILD 240
Qy	241	CDMAPOQLWHSYLTLELNDIVLIGPRKYVDTHNTAEQFLNDPVLIESLPETATNNN 300
Db	241	CDMAPOQLWHSYLTLELNDIVLIGPRKYVDTHNTAEQFLNDPVLIESLPETATNNN 300
Qy	301	PSITSKGNISLDWLEHFKKTDNLRCDSPFRYFVAGNVAFSKWLNKVGWDFEENHWG 360
Db	301	PSITSKGNISLDWLEHFKKTDNLRCDSPFRYFVAGNVAFSKWLNKVGWDFEENHWG 360
Qy	361	GEDVEFGYRLFPAKCFRVIDGGMALHQPPEPKENETEREAGKSITLKVKEKPYIYRK 420
Db	361	GEDVEFGYRLFPAKCFRVIDGGMALHQPPEPKENETEREAGKSITLKVKEKPYIYRK 420
Qy	421	LLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVDLEVCICNDGSTNTLEVI 480
Db	421	LLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVDLEVCICNDGSTNTLEVI 480
Qy	481	NKLYGNPRVRIMSKPNGGSIASASNAAVSFAGYIYGQDSDDDYLEPDAVELCLKEFLKD 540
Db	481	NKLYGNPRVRIMSKPNGGSIASASNAAVSFAGYIYGQDSDDDYLEPDAVELCLKEFLKD 540
Qy	541	KLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHHFRMFTIRAWHLTDGFNEN 600
Db	541	KLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHHFRMFTIRAWHLTDGFNEN 600
Qy	601	TENAVDYDMFLKLSVGEKFLHNLKICYNRVLHGDNTSIKLGIOKKHFFVNVQSLNRQG 660
Db	601	TENAVDYDMFLKLSVGEKFLHNLKICYNRVLHGDNTSIKLGIOKKHFFVNVQSLNRQG 660
Qy	661	INYNYDKFDDDESRRKYIFNKTAIEYQEEIDILKQIKVQRKAKVAISIFYPNTLGLV 720
Db	661	INYNYDKFDDDESRRKYIFNKTAIEYQEEIDILKQIKVQRKAKVAISIFYPNTLGLV 720
Qy	721	KLNNIIEYKNKIFVILLHVDKNHLPDIIKKEILAFYHKHQNILLNDISYTTSNRLIK 780
Db	721	KLNNIIEYKNKIFVILLHVDKNHLPDIIKKEILAFYHKHQNILLNDISYTTSNRLIK 780
Qy	781	TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAMKDYDGMNFSALPHDWEIKIN 840
Db	781	TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAMKDYDGMNFSALPHDWEIKIN 840

Best Local Similarity 86.8%; Pred. No. 5.9e-225;		Matches 844; Conservative 61; Mismatches 60; Indels 7; Gaps 2;	
Qy	1	MNTLSQAIKAYNSNDYELALKLFPSAEYGRKIVFQIKCKEKL---STNS---YVS	53
Db	1	MNTLSQAIKAYNSNDYELALKLFPSAEYGRKIVFQIKCKEKL---STNS---YVS	60
Qy	54	EDKNSVCDSSLDIATQALLSNVKKLTSESEKNSLKNKWSITGKKSENAEIRKVELP	113
Db	61	KEEVNVCDSPLDIATQALLSNVKKLTSESEKNSLKNKWSITGKKSENAEIRKVELP	120
Qy	114	KDFPKDLVLAPLPHVNDFTWYKNNKSLGKIPVKNKIGLSIIPTFNRSILDTIACL	173
Db	121	KDFPKDLVLAPLPHVNDFTWYKNNKSLGKIPVKNKIGLSIIPTFNRSILDTIACL	180
Qy	174	VNOKTNPFEVVDGSGENLTIIVQYEQKLDIKYVRQDYGQYQCAVRNGLRTAKY	233
Db	181	VNOKTHYPFVIVTDGSGEDLSPIIROYENKLDIYVRQDNGFQASAAARNMGLRAKY	240
Qy	234	DFVSILDCDMPAQOLWVHSYITELLEDNDIVLIGPRKYVDTHNITAEQFLNDPYLESLP	293
Db	241	DFIGLLDCDMPAPLWVHSYVAELLEDLTIIGPRKYIDTQHDIPKDFLNNASLESPL	300
Qy	294	ETATNNPSITSGNISLDRLEHFKKTDLNRLCDSPRYFVAGNVAFSKENLKVGFED	353
Db	301	EYKTNVAAKGEVTVSLDRLEHFKKTDLNRLCDSPRYFVAGNVAFSKENLKVGFED	360
Qy	354	EEFNHWGDEVEFGYRLFAGKGFRRVIDGGMALHQPPEKNETERAGKSTIKIVKEK	413
Db	361	EEFNHWGDEVEFGYRLFAGKGFRRVIDGGMALHQPPEKNETERAGKSTIKIVKEK	420
Qy	414	VPIYRKLLPIEDSHIRPLVSIYIPAYNCANYIQRVDSALNQIVDLEVCINDGST	473
Db	421	VPIYRKLLPIEDSHIRPLVSIYIPAYNCANYIQRVDSALNQIVDLEVCINDGST	480
Qy	474	DNTLEVINKLYGNPRVRIMSKPNGGIASASNAAYSFAGYYIGQLSDDDYLEPDAVELC	533
Db	481	DNTLEVINKLYGNPRVRIMSKPNGGIASASNAAYSFAGYYIGQLSDDDYLEPDAVELC	540
Qy	534	LKEFLKDKTLACVYTTNRNVPDGLIANGYNWPFSEKLTAMTAHFMETIRAWHL	593
Db	541	LKEFLKDKTLACVYTTNRNVPDGLIANGYNWPFSEKLTAMTAHFMETIRAWHL	600
Qy	594	TGCFNEINAVDYMFLKSEVKGFKHLNKICYNRVLHGDNTSIKKLGIOKKNHVVVN	653
Db	601	TGCFNEINAVDYMFLKSEVKGFKHLNKICYNRVLHGDNTSIKKLGIOKKNHVVVN	660
Qy	654	QSLNRQGINVYNDKFDLDESRYIPNKTAEYQBEEDMLKDLKIQNKDAKTAVSFYP	713
Db	661	QSLNRQGINVYNDKFDLDESRYIPNKTAEYQBEEDMLKDLKIQNKDAKTAVSFYP	720
Qy	714	NTLGLVKKLNLIYENKNTFVILHVDKNHLPDIIKKEILAFYKHQVNLINNDISYY	773
Db	721	NTLGLVKKLNLIYENKNTFVILHVDKNHLPDIIKKEILAFYKHQVNLINNDISYY	780
Qy	774	TSNRLIKTEAHLNKLNSQLNCEYIIFDNHDSLFVKNDSYAMKYDVGMMFSALTH	833
Db	781	TSNRLIKTEAHLNKLNSQLNCEYIIFDNHDSLFVKNDSYAMKYDVGMMFSALTH	840
Qy	834	DWIEKINHPFKLITFYNDNDRSNVKGASQGMFKYALPHELLTIKEVITSCQS	893
Db	841	DWIEKINHPFKLITFYNDNDRSNVKGASQGMFKYALPHELLTIKEVITSCQS	900
Qy	894	IDSPEYNTEDIWFQFALLILEKKTGHVFNKSTITITYPWPKLQWNTNEQTSKAKGNI	953
Db	901	IDSPEYNTEDIWFQFALLILEKKTGHVFNKSTITITYPWPKLQWNTNEQTSKAKGNI	960
Qy	954	PVNKEFIINSITL 965	
Db	961	PVNKEFIINSITL 972	

RESULT 5

O68389	PRELIMINARY;	PRT;	972 AA.
ID	O68389;		
AC	O68389; 1998 (TrEMBLrel. 07, Created)		
DT	01-AUG-1998 (TrEMBLrel. 07, Last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	Hyaluronan synthase.		
GN	PHAS OR HAS.		
OS	Pasteurella multocida.		
OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;		
OC	Pasteurella.		
OX	NCBI_TaxID=747;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-P-1059, AND ATCC 15742;		
RX	MEDLINE=98192645; PubMed=9525958;		
RA	DeAngelis P.L., Jing W., Drake R.R., Achyuthan A.M.;		
RT	"Identification and molecular cloning of a unique hyaluronan synthase		
RT	from Pasteurella multocida.";		
RL	J. Biol. Chem. 273:8454-8458(1998).		
RP	[2]		
RP	SEQUENCE FROM N.A.		
RA	Fuller T.E., Kennedy M.J., Lowery D.E.;		
RT	"Identification of Pasteurella multocida virulence genes in a		
RT	septicemic mouse model using signature-tagged mutagenesis.";		
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF036004; AAC38318.1; -		
DR	EMBL; AF237926; AAF68412.1; -		
DR	HSP; P36621; 10GQ.		
DR	InterPro: IPR001173; Glycos transf. 2.		
DR	Prfam: PF00535; Glycos transf. 2; 2		
SQ	SEQUENCE 972 AA; 111838 MW; 7D65D024C41ED362 CRC64;		

Query Match 88.2%; Score 4486.5; DB 2; Length 972;
Best Local Similarity 86.6%; Pred. No. 1.4e-224;
Matches 842; Conservative 63; Mismatches 60; Indels 7; Gaps 2;

Qy	1	MNTLSQAIKAYNSNDYELALKLFPSAEYGRKIVFQIKCKEKL---STNS---YVS	53
Db	1	MNTLSQAIKAYNSNDYELALKLFPSAEYGRKIVFQIKCKEKL---STNS---YVS	60
Qy	54	EDKNSVCDSSLDIATQALLSNVKKLTSESEKNSLKNKWSITGKKSENAEIRKVELP	113
Db	61	KEEVNVCDSPLDIATQALLSNVKKLTSESEKNSLKNKWSITGKKSENAEIRKVELP	120
Qy	114	KDFPKDLVLAPLPHVNDFTWYKNNKSLGKIPVKNKIGLSIIPTFNRSILDTIACL	173
Db	121	KDFPKDLVLAPLPHVNDFTWYKNNKSLGKIPVKNKIGLSIIPTFNRSILDTIACL	180
Qy	174	VNOKTNPFEVVDGSGENLTIIVQYEQKLDIKYVRQDYGQYQCAVRNGLRTAKY	233
Db	181	VNOKTHYPFVIVTDGSGEDLSPIIROYENKLDIYVRQDNGFQASAAARNMGLRAKY	240
Qy	234	DFVSILDCDMPAQOLWVHSYITELLEDNDIVLIGPRKYVDTHNITAEQFLNDPYLESLP	293
Db	241	DFIGLLDCDMPAPLWVHSYVAELLEDLTIIGPRKYIDTQHDIPKDFLNNASLESPL	300
Qy	294	ETATNNPSITSGNISLDRLEHFKKTDLNRLCDSPRYFVAGNVAFSKENLKVGFED	353
Db	301	EYKTNVAAKGEVTVSLDRLEHFKKTDLNRLCDSPRYFVAGNVAFSKENLKVGFED	360
Qy	354	EEFNHWGDEVEFGYRLFAGKGFRRVIDGGMALHQPPEKNETERAGKSTIKIVKEK	413
Db	361	EEFNHWGDEVEFGYRLFAGKGFRRVIDGGMALHQPPEKNETERAGKSTIKIVKEK	420
Qy	414	VPIYRKLLPIEDSHIRPLVSIYIPAYNCANYIQRVDSALNQIVDLEVCINDGST	473
Db	421	VPIYRKLLPIEDSHIRPLVSIYIPAYNCANYIQRVDSALNQIVDLEVCINDGST	480
Qy	474	DNTLEVINKLYGNPRVRIMSKPNGGIASASNAAYSFAGYYIGQLSDDDYLEPDAVELC	533
Db	481	DNTLEVINKLYGNPRVRIMSKPNGGIASASNAAYSFAGYYIGQLSDDDYLEPDAVELC	540

```
Qy 534 LKEFLKDKTLCACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHHFRMFTIRAWHL 593
Db 541 LKEFLKDKTLCACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHHFRMFTIRAWHL 600
Qy 594 TDGFENENIENAVDYMFLKSEVGKFKHLNKCYNRVLUHGDNTSIKGLGKIOKKHFFVYVN 653
Db 601 TDGFENIENAVDYMFLKSEVGKFKHLNKCYNRVLUHGDNTSIKGLGKIOKKHFFVYVN 660
Qy 654 QSLNRQGINYYNDFDLDLSRKYIFNKTAEOYOEEMDKLKLIONKDAKIAVISIYP 713
Db 661 QSLNRQGIYYNDFDLDLSRKYIFNKTAEOYOEEDILDKILQNKDAKIAVISIYP 720
Qy 714 NTLNGLVKKLNIIEYNKNKNIIVIIHLVDKNHLLTPDIKEILAFYHKHQNILLNNDISYY 773
Db 721 NTLNGLVKKLNIIEYNKNKNIIVIIHLVDKNHLLTPDIKEILAFYHKHQNILLNNDISYY 780
Qy 774 TSNRLIKTEAHLNINLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDYGMPFSALTH 833
Db 781 TSNRLIKTEAHLNINLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDYGMPFSALTH 840
Qy 834 DWIEKINAPFPFKLIKTYFNNDLRSNMVKGASQGMFKYALPHELLTIKEVITSQS 893
Db 841 DWIEKINAPFPFKLIKTYFNNDLRSNMVKGASQGMFKYALPHELLTIKEVITSQS 900
Qy 894 IDSVPYNTEDIWFQFALLILEKKTGHVFNKTSTLTYPMPWERKLQWNTNEQISAKKGENI 953
Db 901 IDSVPYNTEDIWFQFALLILEKKTGHVFNKTSTLTYPMPWERKLQWNTNEQISAKKGENI 960
Qy 954 PVNKFINSITL 965
Db 961 PVNKFINSITL 972

RESULT 6
Q92IF9 PRELIMINARY; PRT; 604 AA.
AC Q92IF9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Putative two-domain glycosyltransferase.
GN RC0461.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MALISH 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL Science 293:2093-2098(2001).
DR EMBL; AE008609; AAL02999.1; -.
DR InterPro; IPR001173; Glycos_transf_2.
DR Pfam; PF00535; Glycos_transf_2; 2.
KW Transferase; Hypothetical protein; Complete proteome.
SQ SEQUENCE 604 AA; 69121 MW; 514CDD81DC3FB3AB CRC64;

Query Match 5.8%; Score 294; DB 16; Length 604;
Best Local Similarity 24.8%; Pred. No. 1.3e-07;
Matches 124; Conservative 71; Mismatches 213; Indels 92; Gaps 19;

Qy 419 RKLPIEDSHIRPLVSIYIPAYNCANYIQCVDLSALNQTVVDLEVCICNDGSDTN-TL 477
Db 4 RKDKPLNIYH----TLVSIIPVYNGANYMKEAINSAQTYKNEIIIVNDGSKDNGET 59
Qy 478 EVINKLGNPRVRIMSPNGGIASNAAVSAFKAYIGQLDSDDDYLEPDAVE-----LC 533
Db 60 ERVALSYGD--KIRFYFKNGCGGSALNGYIGKNGMQGFYSWLSHDDIYYPNKHIEQV DIL 117
```

```
Qy 534 LKEFLKDKTLCACVY-----TTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHHFRMF 586
Db 118 NKLDNKDFTIYGGYELIDEKGNLSRYIKPDSVLPINKLN---ISLLPLRLGLIHGCSLLM 174
Qy 597 TIRAWHLTDGFENENIENAVDYMFLKSEVGKFKHLNKCYNRVLUHGDNTSIKGLGKIOKK 646
Db 175 PAKYFHEVGIFNEALPTTQDYDLWFKIFRPVPIHFEDESILIKSRFHSQGS--KITSNHN 233
Qy 647 N-----HFVVVNSQLNRQGINYYNDFDLDLSRKYIFNKTAEOYOEEMDKLKL 697
Db 234 ECNVLSFSLHELTEEMIKMEGSPYL-----FLTRATFLSNNT-PYKKACDLANTWA 286
Qy 698 LIQNKDAKIAVISIYPNTLNGLVKKLNNI-IEYNKNKNIIVIIHL-----VDKNHLLTPDIK 752
Db 287 KOVLNDTRKISVIIPVYNINNAIEAKSVLIQTHKNFEILLIDCGSTDIDSELTAICKRD 346
Qy 753 -ILAFYHKHQNILLNNDISYITSNRLIKTEAHLNINLSQNLNCEYIIFDNHDSLFV 811
Db 347 KRKYFHK-----KNEGPAARN-LGIKNAIGKYIAFLDSDDLFY 385
Qy 812 KNDYAYMKKYDVGWGNFSALTHDWIEKINAHPPFKLIKTYFNNDLRSNMVKGASQGMF 871
Db 386 K-DKIEIQLKWEENNF-IFSHTSYHKINE-----KCKYIESVHSGLF 426
Qy 872 MKYALPHELLTIKEVITSC 891
Db 427 SGNVFP-----QVIQTC 438

RESULT 7
Q8Y828 PRELIMINARY; PRT; 327 AA.
AC Q8Y828;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein lmo1090.
GN LMO1090.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-E / SEROVAR 1/2A;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstedt G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlund J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL591977; CAC99168.1; -.
DR Listlist; LMO1090; -.
DR InterPro; IPR001173; Glycos_transf_2.
DR Pfam; PF00535; Glycos_transf_2; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 327 AA; 37843 MW; F7480656043EC2FC CRC64;

Query Match 5.7%; Score 292.5; DB 16; Length 327;
Best Local Similarity 27.5%; Pred. No. 7.7e-08;
Matches 98; Conservative 44; Mismatches 133; Indels 81; Gaps 9;

Qy 433 PLVSIYIPAYNCANYIQCVDLSALNQTVVDLEVCICNDGSDTNTLEVINKLGNPRVRI 492
Db 3 PLVSIYIPVYNNKRYKRCLDVLEQTVHNLEVIIVVNDGATNSAKVIKISDN--RIRY 60
```

Db 284 IKRLDYCYLSVDNKLINILYFLR-----DKNN-----SYFNKKQ 319

RESULT 9

Q9ALS8 PRELIMINARY; PRT; 350 AA.

AC Q9ALS8; 01-JUN-2001 (Tremblrel. 17, Created)

DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)

DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)

DE Hypothetical 41.6 kDa protein.

OS Campylobacter jejuni.

OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;

OC Campylobacter.

OX NCBI_TaxID=197;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NCTC 11828;

RA Oldfield N.J., Millar L.A., Ketley J.M.;

RT "Gene content polymorphisms in Campylobacter jejuni LOS biosynthesis clusters";

RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF343914; AAK12955.1; ~

DR InterPro; IPR001173; Glycos_transf_2.

DR Pfam; PF00535; Glycos_transf_2; 1.

KW Hypothetical protein.

SQ SEQUENCE 350 AA; 41616 MW; FB19017F89326E9E CRC64;

Query Match 5.6%; Score 287.5; DB 2; Length 350;

Best Local Similarity 25.5%; Pred. No. 1.5e-07;

Matches 105; Conservative 66; Mismatches 129; Indels 111; Gaps 16;

QY 433 PLVSIYIPAYNCANQIQRCDVSALNQTVVDLEVCICNDGSDTNTLEVINKLYGNPNRVRI 492

Db 5 PLVSIIPAYNCANQIQRCDVSALNQTVVDLEVCICNDGSDTNTLEVINKLYGNPNRVRI 64

QY 493 MSKPNGGIAS-ASNAVSEAKGYIGQLSDSDYLDPAVELCLKEFLKDKTKLACVYT--- 548

Db 65 YSINHGTGIPSVVKNYGLKAKGFEFLILSDDDMTIEYFLEKGIKVFQDPVDILYPIKF 124

QY 549 --TNRNVNPDGSLIANGYNWPEFSREKLTAMTAHFRMFTTRAWHLTDCGFENIENAVD 606

Db 125 MFSNNYKIIGIYNNSLNIDSVNLCATNKIISGR-----DAFRNIYNNKLI 172

QY 607 YDMFLKSEVGKFKHLNKCINRVLHGDNTSIRKLGIOKK-----NHFFVVAQSLNRQG 660

Db 173 GFPEYK-KTIDKIINFNEESFN---GDEYSFRELHQAQKIAFIIDTEFVYIN--FNQES 225

QY 661 I-----NYNYDKFDLDESRYIFNKTAEOEEMDMKLDKLIQNKDKAKIA 707

Db 226 ITKKIGVHHWDTKWFNLEKL-----AQHNVEK-----KLK-KINKIR 265

QY 708 VSIFYPNTLNGLVKKLNIIYNNKNIIVLHVDKNHLPDIIKKEILAFYHKKHVNILLN 767

Db 266 YSIYV-----ELCIKFNKTEYL-----PSQNEKNIL- 292

QY 768 NDISYVTSNRLLKTEAHLNINLKLSQ-LNLNC-----EYIFDNHSLFVK 812

Db 293 -----NKILEKNHLSRLINSIFDFLFYCKDEKGYIKFYKTYFYK 335

RESULT 10

Q8XN37 PRELIMINARY; PRT; 326 AA.

AC Q8XN37; 01-MAR-2002 (Tremblrel. 20, Created)

DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

DE Capsular polysaccharide biosynthesis protein.

GN CP50501.

OS Clostridium perfringens.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;

QY 493 MSKPNGGIASNAVSEAKGYIGQLSDSDYLDPAVELCLKEFLKDKTKLACVYTNNR 552

Db 61 FEKENGQATARNFGLDVAGDYIVMVDSDYISKNIVETCL-DTVQKTNADLVFTSYN 119

QY 553 VNPDGSS-----LIANGYNWPEFSREKLTAMTAHFRMFTTRAWHLTDGPF--- 597

Db 120 VNQEGKQYIKRDKGIKVLDAGPTPNKFKYQADLWKG-----SRFPVGYWYEDLGIIPV 173

QY 598 -----NENTENAVDY---DMFLKSEVGKFKHLNKCINRVLHGDNTSIRKLGIOKKN 647

Db 174 VTLKAKNPVKIQDALYYITDRADSQSIQVDFHLDVVI--MLENVETELKLGII--- 227

QY 648 HFVVVQSLNRQGINYYNDRFDDLESRYIFNKTAEOEEMDMKLDL-----KLIQNKDA 704

Db 228 -----YEEKDQLAYLYIEHLIYRLV 249

QY 705 KIAVSIFYPNTLNGLVKKLNIIYNNKNIIVLHVDKNHLPDIIKKEILAFYHKKH 760

Db 250 RKAIVITNKQERKKLIKISTIOEQFPNWSYPYQAGGKLTATLKKRALWLYLHH 305

RESULT 8

Q9CLR9 PRELIMINARY; PRT; 337 AA.

AC Q9CLR9; 01-JUN-2001 (Tremblrel. 17, Created)

DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE Hypothetical protein PM1140.

GN PM1140.

OS Pasteurella multocida.

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

OC Pasteurella.

OX NCBI_TaxID=747;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PM70;

RX MEDLINE=21145866; PubMed=11248100;

RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;

RT "Complete genomic sequence of Pasteurella multocida PM70.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).

DR EMBL; AE006155; AAK03224.1; ~

DR InterPro; IPR001173; Glycos_transf_2.

DR Pfam; PF00535; Glycos_transf_2; 1.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 337 AA; 39267 MW; 8639BCFB5F700DB4 CRC64;

Query Match 5.7%; Score 292; DB 16; Length 337;

Best Local Similarity 27.0%; Pred. No. 8.5e-08;

Matches 96; Conservative 67; Mismatches 124; Indels 68; Gaps 16;

QY 433 PLVSIYIPAYNCANQIQRCDVSALNQTVVDLEVCICNDGSDTNTLEVINKLYGNPNRVRI 492

Db 7 PLVSLICAYNADRIIECDAILNQYKNLEIVVVDGSDTNTLKLHYPAGKDPRIKI 66

QY 493 M-SKPNGGIASNAVSEAKGYIGQLSDSDYLDPAVELCLKEFLKDKTKLACV---YTT 549

Db 67 INNEENKGFASLNIGIASINGDYLDARTDADITKPEWIEKILGYMLSHPOIIAMGSYLT 126

QY 550 NRVNPDGSLIANGYNWPE-FSREKLTAMTAH-----FRMFTTRAW 591

Db 127 ILSEGDGNSLNAYEHEGDEWRNPLSHREIVEAMLFNPIHNSMIVKSTVFRHGLR-- 184

QY 592 HLTDGFENIENAVDYMFLKSEVGKF-KHLNKCINRVLHGDNTSIRKLGIOKKHNFV 650

Db 185 -----FDPAYQHTEDYQFWEVSRGELANYPESLVYVR-LH--NTQTSSLHKNYQ--- 233

QY 651 VVQNSLRQGINYYNDRFDDLESK-----YIFNKTAEOEEMDMKLDKLIQNKDKAKI 706

Db 234 LMAKKIRKRAINY----LQDLGVIRHGEDIFFHDIETIQAE---LASLSLDN---CI 283

QY 707 AVSIFYPNTLNGLVKKLNIIYNNKNIIVLHVDKNHLPDIIKKEILAFYHKKH 761


```

RN      SEQUENCE FROM N.A.
RP      STRAIN=TIGR4;
RC      MEDLINE=21357209; PubMed=11463916;
RA      Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA      Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA      Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA      Unayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA      McInnis J.E., Khouri H., Wolf A.M., Uterback T.R., Hansen C.L.,
RA      McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA      Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C., C.M.;
RA      Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT      "Complete genome sequence of a virulent isolate of Streptococcus
RT      pneumoniae";
RL      Science 293:498-506(2001).
DR      EMBL; AE007470; AAK75845.1;
DR      TIGR; SPI771;
DR      InterPro; IPR001173; Glycos_transf_2.
DR      InterPro; IPR002495; GT_8.
DR      Pfam; PF00535; Glycos_transf_2; 1.
DR      Pfam; PF01501; Glyco_transf_8; 1.
DR      Transferrase; Complete proteome.
KW      Transferrase; Complete proteome.
SQ      SEQUENCE 696 AA; 81336 MW; ADCA9C01AAB32C7 CRC64;

Query Match      5.3%; Score 268.5; DB 16; Length 696;
Best Local Similarity 20.5%; Pred. No. 3.3e-06;
Matches 101; Conservative 96; Mismatches 181; Indels 115; Gaps 17;

QY 435 VSIYIPAYNCANYIQCVCDSALNQTVWDLVCICNDGSTDNTFLVINKLYGNPNVRIMS 494
DB 7 ITIVFVNVNENYLRKCLDSITQYKNIIVVNDGSTDASGEICKEPSEMDHRILEYE 66
QY 495 KPNGGIASASNAASFAGYIYGQDSDYLEPDADVELCLK---FELKDKTLACVYTTNR 551
DB 67 QENAGLSAARNGLNMSGNYTFVDSDDWIEQDYETLYKKIVEYQADIAVGNYSFNE 126
QY 552 N-----VNPDSGLIANGYNWPEFSREKLTMTAMIAHFMFTIRAWHL-- 593
DB 127 SEGMYFHILGDSYKVEKVDNVSIFENLYE---TQEMKSFALISAWGLKARLPEQLR 182
QY 594 -----TDFGNENIENAVDYDMFLKSEVGKFKHLNKICY-NRVLUHGD----- 634
DB 183 FDIGKLGEDGY-----LNQKVL-LSE--KVIYLNKSLIYAIRIRKGLSLRWTEKWMH 232
QY 635 -----NTSKKLGIGQKKNHFVNVQSLNRQGNINYNDKFDLDESRYENKTAE 685
DB 233 ALVDMASERITLLANNGYPLEKHLAVYRQMLEVSLAN---GOAGSLSDATY-----KE 283
QY 686 YQEMDMKDLKLIQNKDAKIAVSIFYPNTPLNGLVKKNLNNIIEYKNIFVILHVD-KNH 744
DB 284 FEMKQRLNLQSRQESSEKKAIVLAANYGYVDQVLTITKISYCHNRSIRFYLIHSDFNE 343
QY 745 LTPDIKKEILAF-----YHKQVNLNNDISYTSNRLIKTEAHLNINKLSQNL 796
DB 344 WIKOLNKRLEKFOSEIINCRTVTEQIS-CYKSDISYTVFLRYTADF---VOEDKKALYL 398
QY 797 NCEYIIIFDNHDSLFVKN-DSY-----AYMKKYDVGMM 827
DB 399 CDCULVYTKNLDLFDATQDYPLAAVRDFGGRAYGQEIFNAGVLLYNNFAWKENTQK 458
QY 828 FSALTDHWIEKIN 840
DB 459 LIDVTNEWHDKVD 471

RESULT 13
Q8XN34 ID Q8XN34 PRELIMINARY; PRT; 348 AA.
AC Q8XN34;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Spore coat polysaccharide biosynthesis protein.

```

```

GN      CPE0504.
OS      Clostridium perfringens.
OC      Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC      Clostridiales; Clostridiaceae; Clostridium.
OX      NCBI_TaxID=1502;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=13 / TYPE A;
RX      PubMed=11792842;
RA      Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA      Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT      "Complete genome sequence of Clostridium perfringens, an anaerobic
RT      flesh-eater";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR      EMBL; AP003187; BAB80210.1;
DR      InterPro; IPR001173; Glycos_transf_2.
DR      Pfam; PF00535; Glycos_transf_2; 1.
KW      Complete proteome.
SQ      SEQUENCE 348 AA; 41769 MW; AB46C19FA8E52C93 CRC64;

Query Match      5.2%; Score 265.5; DB 16; Length 348;
Best Local Similarity 23.3%; Pred. No. 2.1e-06;
Matches 109; Conservative 68; Mismatches 132; Indels 159; Gaps 18;

QY 435 VSIYIPAYNCANYIQCVCDSALNQTVWDLVCICNDGSTDNTFLVINKLYGNPNVRIMS 494
DB 6 VSIIVPAYNIESYVERCLNSLINQTYENIEIIVDGDSTNTLVKISDMARKDKRIKIVLE 65
QY 495 KPNGGIASASNAASFAGYIYGQDSDYLEPDADVELCLKFEKDKTLACVYTTNRNRYN 554
DB 66 QENRGSSEARKKGYEMSTGEFILFVDGDDWIRNDIEVLE-----YS-----N 109
QY 555 PDGSLIANGYNWPEFSREKLTMTAMIAHFMFTIRAWHLTDGFNENIENAVDYDMFLKLS 614
DB 110 DDIDIVAFGYN-----EIFNEN----- 126
QY 615 EVGKFKHLNKICYNRVLHGDNTSIKKLGQKKNHFVNVQSLNRQGNINYNDKFDLDE 674
DB 127 EIQK-----NSLYDK-----NNEKIELRDTQFLREILTNNISIN-----INWF---I 167
QY 675 SRKYIFNKTAEYQEMDMKDLKLIQNKDAKIAVSIFYPNTPLNGLVKKNLNNIIEYKNIF 734
DB 168 RKFFIDKNVNVFPKMSYAEADALLISLAAK-----EPNVI 203
QY 735 VIILHVDKNHLPDIKKEILAFYHKHQ-----VNILLNNDISYTSNRLIKTEAHLN 786
DB 204 VI-----KEKLYFYFKRENSTISISKILETKAMFIKNILIE----- 243
QY 787 NINKLSQNLN---CEYI---IFDNHDSLFVKNDSYAYMKKYDVGMMFNSALTDHWIEKIN 840
DB 244 --NKLNIYEEFEYCYVIHNNWYFRYRIIFYGNQYS-KKLFEIWNFNFI-----SIR 293
QY 841 AHPFPFKLIKLT---YFNDNDRSMNVKVGASQGMFKYALPHELLTIK 885
DB 294 NNKYFKKDEKNLDCYWRFFYLNAIN-----NYILGKILALIIR 331

RESULT 14
Q8YSL7 ID Q8YSL7 PRELIMINARY; PRT; 323 AA.
AC Q8YSL7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein Alr3067.
GN Anabaena sp. (strain PCC 7120).
OS Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,

```


RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpō S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.,
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003591; BAB74766.1;
DR InterPro; IPR001173; Glycos_transf_2;
DR Pfam; PF00535; Glycos_transf_2; 1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 323 AA; 36955 MW; 0AA1437F9EC42D89 CRC64;

Query Match 5.2%; Score 264.5; DB 16; Length 323;
Best Local Similarity 25.4%; Pred. No. 2.1e-06;
Matches 82; Conservative 64; Mismatches 116; Indels 61; Gaps 12;

QY 432 IPLVSIYPAYNCANYIQRCDVSALNQTVVLEVCICNDGSDTNTLEVINKLYGNPRVR 491
DB 10 VPIISVIIIPVANGKTIETIASVQHQTFLDIETIIVINDGSTDTNPFELVRNQDN--RLK 67

QY 492 INSKNGGTASASNAVSFAKGYIIGDSDDDYLEPDAVEL---CLKEFLKDKTKLACVYT 548
DB 68 IFSYENGGLPVARNRGITHAVGQFIATDADDLWTFDKLEQFAALQEQY-PEAGLAYSWT 126

QY 549 TNRNVNPDGS-----LIANGYNWPEFSREKLTMTAMIAHHRFMETI 588
DB 127 YKPFANEADSYADESNSFAGDYVYAEILLKFNQNGSN-PLIRRAIDSVGL----- 176

QY 589 RAWHLTDFGNENIENAVDYDMFLKISEVGKFKHLNKK--ICYNRVLHGDNTSIKKLGIOKK 646
DB 177 -----FDPFLKCEDWDFYLRLLAAKQWQFALVKKAAQIIYRQ---SPTAMTSKLDVMEK 225

QY 647 NHFVVVNSLNQGINYYNDKFDLDESRYIFNKTAEE--YQDEMDMLKDLKLQNKDA 704
DB 226 YSSIVIERAFNAAPLOLHLKK-----OSLAWVYKFTAOCLCKYNSHKLADIKLAA-KRL 279

QY 705 KIAVSIFYPNTL----NGLVKRL 723
DB 280 KNAITLPKNLLEDYTHGLIRKL 302

RESULT 15
Q9K6L6 PRELIMINARY; PRT; 303 AA.
AC Q9K6L6;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE Glycosyltransferase.
GN BH3713;
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-125 / JCM 9153;
RA MEDLINE-20512582; PubMed-11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirana C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001519; BAB07432.1;
DR InterPro; IPR001173; Glycos_transf_2;
DR Pfam; PF00535; Glycos_transf_2; 1;
KW Transferase; Complete proteome.
SQ SEQUENCE 303 AA; 35235 MW; C9F5471D29CD6076 CRC64;

Query Match 5.1%; Score 261.5; DB 16; Length 303;
Best Local Similarity 26.5%; Pred. No. 2.8e-06;

Matches 90; Conservative 60; Mismatches 95; Indels 95; Gaps 15;
QY 435 VSLIPIPAYNCANYIQRCDVSALNQTVVLEVCICNDGSDTNTLEVINKLYGNPRVRIMS 494
DB 3 VSIITPHRAKLLKRALESTLQTYKNIEVIVVSDGSTNT-DIVNDEYKRDSTRVNFIS 61

QY 495 -KPNGGISASNAASVFAKGYIIGDSDDDYLEPDAVELCLKEFLKDKTKLACVYT----- 548
DB 62 YHPAKGNGYARNITGINKNAGEFIATLDDDDDEWMPDKLELQIKFENQANVGLVYTGVEII 121

QY 549 ----TNRNVNPDGSLIANGYNWPE-----FSREKL-----TTAMIAHHRFMFTIRAWHLTD 595
DB 122 YNFENKRN-----KYISLPKKTGNLSKEILVANCIGTSSV-----MYRKNLITE 166

QY 596 --GFENIENAVDYDMFLKLSKSEVGKFKHLNKCYNRVLHGDNTSIRKLGIOKKNHFVVVN 653
DB 167 CGMFEKLRARQYDILWI-----RVC-----QKTLVGVVN 196

QY 654 QSLNRQGINYYNDKFDLDESRYI-----IFNKTAEE-----YQDEMDMLK 694
DB 197 KPL----VRYNYNTNKOISDDIKKYESAIEYIDNKYVDLYSKVSEIRKKHRSMTMLI 252

QY 695 DKLIONKDAKIA-----VSIFYPNTLNGLVKKLNIIIEY 729
DB 253 VNKALRNQSPKVARAYLKNLSFLKRPRTLTAIMYMLSLFKY 292

Search completed: January 4, 2003, 02:26:39
Job time : 127 secs

THIS PAGE BLANK (USPTO)